MUHAMMAD NURIL HUDA

nurilhuda3333@gmail.com LinkedIn GitHub

AI Engineer | Data Scientist | Python, SQL, LLM Systems

+6281334374014 Indonesia

SUMMARY

AI Engineer and Data Scientist with a Master's degree in Computer Science and hands-on experience building practical data and AI systems using **Python, SQL, FastAPI, PySpark, Docker, and cloud platforms**. Strong fit for product-driven AI roles that require turning messy data into simple, useful, and actionable features.

Experienced in building data pipelines, forecasting workflows, backend APIs, RAG-based applications, and LLM-assisted systems. Previously worked with large transportation datasets at Data Glacier, integrating taxi, flight, and weather data for forecasting and operational insight generation. Strong interest in building fast, business-oriented AI features such as anomaly detection, automated insights, benchmarking, and LLM-generated explanations.

CORE EXPERTISE

Data-Driven Features: forecasting, anomaly-pattern analysis, benchmarking logic, automated insights
Backend & Data Systems: Python, SQL, FastAPI, REST APIs, batch data processing, Docker
Practical AI: LLM APIs, RAG pipelines, prompt design, structured explanations, AI-assisted workflows
Business Translation: turning ambiguous requirements into technical workflows, metrics, and usable outputs

EXPERIENCE

AI Evaluation Consultant / STEM AI Trainer Oct 2024 – May 2026
Alignerr, United States

- Evaluate AI-generated solutions for STEM, coding, machine learning, and data science reasoning tasks, focusing on correctness, instruction-following, and practical usefulness.
- Review Python and code-related outputs by identifying logical errors, edge-case failures, inefficient approaches, and incomplete implementations.
- Produce structured technical feedback and written rationales to support high-quality LLM training data and model evaluation workflows.
- Serve as a reviewer/auditor for data-labeler outputs, checking rubric adherence, factual accuracy, reasoning quality, and annotation consistency.
- Apply independent judgment in ambiguous tasks, balancing technical correctness with clarity, usefulness, and user-facing quality.

AI Trainer / LLM Evaluator May 2024 – Oct 2025
Outlier, United States

- Reviewed and annotated AI-generated responses for RLHF and SFT tasks across general, technical, and computer science-related domains.
- Assessed outputs across instruction-following, truthfulness, completeness, clarity, reasoning quality, and overall usefulness.
- Supported AI model improvement through structured evaluation, written feedback, and quality assurance workflows.

- Previously worked on voice-related AI training tasks, strengthening attention to data quality, consistency, and user-facing output standards.

Data Scientist – Internship Jan 2024 – Jul 2024
Data Glacier, India

- Developed data pipelines integrating taxi, flight, and weather datasets using **PySpark**, **Python**, and **Google Cloud Platform**.
- Built time-series forecasting models using **SARIMA** and **LSTM** to predict demand patterns, improving prediction accuracy by 15
- Optimized data processing workflows, reducing runtime by 90
- Worked with large, real-world transportation datasets to identify patterns, support operational recommendations, and translate analytical outputs into business insights.
- Delivered analytical reports and visualizations to communicate findings clearly to stakeholders.

Data Analyst – Apprenticeship Jul 2023 – Dec 2023
GoTo Impact Foundation, Indonesia

- Designed and maintained dashboards for student performance tracking and program reporting.
- Defined learning metrics and automated reporting workflows to support data-driven program evaluation.
- Translated analytical findings into clear dashboards and concise reports for non-technical stakeholders.

AI Engineer – Apprenticeship Aug 2021 – Feb 2022
Orbit Future Academy, Indonesia

- Led a small team to develop a deepfake detection system using CNN-based architectures, including MTCNN and InceptionResNetV1.
- Built modular preprocessing scripts and model evaluation workflows for image classification tasks.
- Coordinated team responsibilities across data preparation, model training, evaluation, and final project delivery.

SELECTED PROJECTS

Hospital Chatbot Appointment System

- Built a backend-driven medical chatbot appointment system using **FastAPI**, **Docker**, relational database workflows, caching, and a **LangChain-based RAG pipeline**.
- Implemented document retrieval, prompt orchestration, multi-intent handling, and automated doctor assignment logic.
- Designed the system to transform a manual appointment workflow into an AI-assisted service with structured backend logic and user-facing responses.

DeciSense – Open-Source AI Assistant for Data Science Workflows

- Developing a Python-based AI assistant project focused on dataset intake, validation, profiling, model recommendation, and AI-assisted data science workflows.
- Designed modular components for tabular data loading, validation, task inference, profiling, planning, modeling, evaluation, and reporting.
- Focused on practical AI automation: using simple, inspectable logic first, then adding LLM assistance for explanation, recommendation, and workflow orchestration.

Transportation Demand Forecasting Pipeline

- Built forecasting workflows on transportation-related datasets by combining taxi, flight, and weather data.
- Applied time-series modeling and data processing optimization to support demand pattern analysis and operational decision-making.

EDUCATION

Master of Computer Science (M.Cs.) Aug 2023 – Jan 2025
Universitas Gadjah Mada, Yogyakarta, Indonesia
GPA: 3.79/4.00
Thesis: Early detection of diabetic foot ulcers using thermogram and temperature data integration, achieving 97.06% accuracy.

Bachelor of Computer Science (S.Kom.) Aug 2018 – Sep 2022
Universitas Muhammadiyah Malang, Indonesia
GPA: 3.86/4.00

PUBLICATIONS & AUTHORSHIP

Optimization of Plantar Foot Thermogram for Diabetic Foot Ulceration Early Detection: An Image Enhancement Approach
SINTECHCOM Journal, Feb 2025
DOI: 10.59190/stc.v5i2.273

Java Itu Mudah
DIVA Press, Dec 2020
ISBN: 9786023918980

SKILLS

Programming & Backend: Python, SQL, FastAPI, REST APIs, Git, Docker, basic backend architecture
Data Engineering: PySpark, ETL pipelines, batch processing, data cleaning, schema design, workflow optimization
Analytics & Modeling: time-series forecasting, statistical analysis, threshold-based logic, clustering, classification, regression, model evaluation
Practical AI & LLMs: LLM APIs, prompt engineering, RAG pipelines, structured output evaluation, automated insight generation, LLM evaluation
Cloud & Deployment: cloud platforms, BigQuery, Compute Engine, Vertex AI exposure, Streamlit, API deployment concepts
Product & Collaboration: business problem framing, stakeholder reporting, technical documentation, asynchronous collaboration, fast iteration