MOHAMMAD ARKANANTA RADITHYA TARATUGANG

DATA & MACHINE LEARNING · PYTHON, SQL, PYTORCH

Jakarta, Indonesia | +62 822-5477-2727 | radithyarkaa@gmail.com | linkedin.com/in/arkananta-radithya | github.com/radithyaarka | www.kaggle.com/arkaradithya | Website: radithyarka.netlify.app | Portfolio (Deck PDF)

Available for full-time work, open to relocation and remote arrangements.

SUMMARY

Data analyst and machine-learning practitioner with a first-author paper presented at ICAICTA 2026: 82.10% accuracy / 79.10% F1 on 294 multimodal tweet-image pairs, with 13 false negatives. Built the streaming and storage layer of a Kafka-MinIO-Spark recommendation pipeline and modelled recommendations in Spark. S.Kom in Information Technology, Institut Teknologi Sepuluh Nopember.

TECHNICAL SKILLS

Languages & Query: Python, JavaScript, TypeScript, SQL (PostgreSQL, MySQL), HTML/CSS, Java, C++.

Data & ML: Pandas, NumPy, Scikit-learn, PyTorch, Machine Learning, Deep Learning, Large Language Models (LLM), Retrieval-Augmented Generation (RAG), Exploratory Data Analysis (EDA), SHAP (XAI), Natural Language Processing (NLP), Computer Vision, Microsoft Excel (Pivot Tables, VLOOKUP).

Engineering & Tools: Apache Kafka, Apache Spark, MinIO (Data Lakehouse), Trino (SQL Query Engine), Docker, Docker Compose, Git, Linux/Bash, FastAPI, Spring Boot, Streamlit, React.js, Next.js, Vue.js (Vue 3/Vite), Node.js (Express), TailwindCSS, RESTful APIs, Object-Oriented Programming (OOP), Figma (UI/UX Design), Postman, Firebase.

EDUCATION

Institut Teknologi Sepuluh Nopember (ITS), Surabaya **2022 - 2026**

S.Kom (Information Technology) - GPA: 3.62 / 4.00

- **Relevant Coursework:** Artificial Intelligence & Machine Learning, Big Data & Data Lakehouse, Database Systems & SQL, Data Structures & Algorithms, Web Programming.

PUBLICATION

The 13th International Conference on Advanced Informatics: Concepts, Theory and Applications (ICAICTA 2026) **Presented 20 October 2026**

First Author · Paper built on the CelestX project

- *Detection of Potential Depressive Disorders on 'X' Social Media Based on User-Level Analysis Using Multi-Modal Transformers Architecture*

WORK EXPERIENCE

Data & Machine Learning Engineer **Aug 2026 - Present**

JokiProyek - Remote

- Handle whatever each client asks for, end to end: write the scope of work and the delivery estimate first, then build it.
- Cover both sides of the request: data work (scraping, cleaning, and pipelines in Pandas and Scikit-learn) and AI work (retrieval-augmented (RAG) chatbots, trained models in PyTorch).
- Ship it as a working tool (FastAPI, Streamlit, or React) and maintain existing client codebases through debugging, performance tuning, and bug fixing, with weekly progress reports to the internal PIC.

Information Technology Intern **Dec 2024 - Feb 2025**

PT. Badak NGL - Bontang, East Kalimantan

- Built an RSA login flow prototype for 3 internal systems (e-Procurement, e-PNK, e-Contract) that share one login form; encrypted credentials with the public key along the login path, with database storage left to the implementation.
- Tested the flow with dummy data, then handed the source code to the IT team and presented and demoed it for implementation.

Teaching Assistant **Feb 2023 - Jun 2025**

Institut Teknologi Sepuluh Nopember (ITS) - Surabaya, Indonesia

- Mentored 80+ students per semester across 4 core courses (Big Data & Data Lakehouse, Web Programming, Mobile Programming, Data Structures & OOP) over 4 semesters.
- Co-authored 5 lab modules and graded 10+ final project teams per semester against code architecture, security, responsiveness, and usability standards.
- Taught Apache Kafka (real-time streaming) and MinIO (data lakehouse) labs and graded end-to-end data analysis final projects; held weekly consultations on programming logic and algorithms.

PROJECTS

CelestX: Multimodal AI Mental Health Observatory github.com/radithyaarka/celestx

Solo Project: model architecture, backend, frontend, XAI

- Trained a multimodal Transformer (IndoBERTweet + ViT) in PyTorch to detect DSM-5 depression indicators, reaching 82.10% accuracy / 79.10% F1 with 13 false negatives on 294 tweet-image pairs (5-fold cross-validation); the text modality was trained on 7,991 Indonesian tweets.
- Built a Chrome Extension, React.js dashboard, and FastAPI backend for real-time monitoring; predictions were explained via SHAP values and attention maps and discussed with a licensed clinical psychologist, with usability scored 80/100 (SUS, "good").

Big Data Movie Recommender System github.com/radithyaarka/fp-big-data

Data Engineer & Backend Developer (team of 5)

- Built the streaming and storage layer of a recommendation pipeline: a Kafka Producer-Consumer feeding a MinIO data lakehouse (raw JSON and Parquet).
- Modelled recommendations in Apache Spark (CountVectorizer + cosine similarity); results were queried through Trino SQL by a teammate and served in a Streamlit interface.

BoothMe: AI Photo Booth github.com/sgraa/boothme

UI/UX Designer & Full-Stack Web Developer (team of 3)

- Built a 3-component system (Vue 3 client with MediaPipe face tracking, Node.js/Express API, Stable Diffusion 1.5 service) in a team of 3, with the UI/UX designed in Figma.
- Streamed high-resolution base64 images from the client to the diffusion service through the Node.js API, orchestrated with Docker Compose.

F1 Race Strategy: Pit-Stop Cost and the Undercut kaggle.com/code/arkaradithya/f1-race-strategy

Data Scientist

- Analysed 366,508 laps in the 2011-2026 Formula 1 window (Ergast archive) to measure pit loss at every circuit with bootstrap confidence intervals and test when undercuts actually convert; the standard pit-loss rule performs at chance level, while pre-stop gap and tyre age predict the outcome out of sample.
- Built a tested Python and SQL pipeline (Parquet/SQLite storage, independent SQL cross-checks) and modelled pit-stop timing one lap ahead with gradient boosting versus a logistic baseline, reporting permutation importance and the precision/recall trade-off.

Retail Demand Forecasting kaggle.com/code/arkaradithya/retail-forecasting

Data Scientist

- Forecasted daily sales on the million-row Rossmann dataset with PySpark Random Forest and train-only target encoding, reaching MAE $886.13 / RMSE $1,256.03 / R^2 0.8370 on a strictly later June-July 2015 holdout (58,611 rows).
- Quantified promotions at a +38.8% average sales lift (from $5,929 to $8,228 per ordinary day) across 844,392 store open-days, answering three business questions on promotion-driven demand.

ORGANIZATIONAL & COMMITTEE EXPERIENCE

Media & Information Staff, promoted to Vice Head of Department **Feb 2024 - Feb 2026**

HMIT ITS - Media Information Department - Surabaya

- Added roughly 3,000 followers to the department's social media across 2 management periods through publication strategy and visual communication.
- Led a team of 11 (8 designers and 3 coordinators) after being promoted to Vice Head of Department: creative direction, publication deadlines, and digital asset QA.
- Co-developed the organization's brand guidelines and applied them across digital channels and print materials.

Head of Design & Documentation **July 2024 - September 2024**

OKKBK IT 2024 (IT Dept. Freshman Orientation Event) - Surabaya

- Led the design and documentation division of OKKBK IT 2024: set the event's visual identity, typography, and digital publication assets.
- Ran the division's workflow, daily task delegation, and QC for all social media publication assets and event documentation.

LANGUAGES

Indonesian (Native), English (IELTS 7.0/9.0).