



Arya Nathan Bara

Graduate Computer Science Student · Portfolio

Jakarta, Indonesia | arya.anb.sky@gmail.com | +62 878-8708-8282 | linkedin.com/in/arya-nathan-bara

PROFILE

Computer Science master's student (Data Science) who turns data into business decisions — end to end across SQL, data warehousing, business-intelligence dashboards, and structured problem-solving. I've designed a star-schema warehouse, shipped an interactive dashboard of sales, delivery, and customer-satisfaction KPIs for non-technical stakeholders, and built a GenAI text-to-SQL system with rigorous statistical evaluation. Alongside the analytics, I've led a 30+ staff student organization of 500+ members and closed corporate sponsorships, so I'm comfortable presenting findings and coordinating across teams. Author on two IEEE conference papers, first author on one.

EDUCATION

Master of Computer Science (Data Science stream) · BINUS Graduate Program Feb 2026 – Present
Jakarta, Indonesia • Expected Feb 2027 • GPA 4.00 / 4.00

- Thesis: a late-fusion transformer framework (IndoBERT, XLM-RoBERTa) combined with psycholinguistic features and SHAP-based interpretability to detect workplace stress in Indonesian text.
- Focus coursework: Business Intelligence & Analytics, IT Risk Management & Audit, Deep Learning, Advanced NLP, Data Warehousing.

Bachelor of Computer Science (Minor in Digital Transformation) · BINUS University Sep 2022 – Feb 2026
Jakarta, Indonesia • Completed via fast-track program • GPA 3.78 / 4.00

- Exchange semester, Artificial Intelligence at Kyung Hee University, South Korea (Feb – Jul 2025); awarded KRW 200,000 for a self-written cultural-exchange grant proposal.

WORK EXPERIENCE

Associate System Analyst · Bina Nusantara IT Division Mar 2023 – Mar 2024

- Partnered with cross-functional teams to translate business requirements into shipped system changes.
- Wrote and optimized SQL queries for internal reporting and resolved recurring user data-quality issues.

ORGANIZATIONAL EXPERIENCE

General President · Modelling Club, BINUS University Feb 2024 – Feb 2025

- Directed a 30+ staff organization serving 500+ members across six campus regions, delivering 30+ programs on schedule.
- Closed 100M+ IDR in corporate sponsorships (Paragon Corp, TikTok Shop, Tokopedia) by building and pitching data-backed partnership proposals.
- Grew recruitment 3.3x (70 → 230 registrants) and raised sign-up fees 133% (75k → 175k IDR, ~40M IDR revenue) via a rebrand and recruitment-funnel analysis.
- Launched an in-house talent agency connecting members to paid modelling and other work.

First Year Program Leader (Binusian 2027) · BINUS University Mar 2023 – Aug 2024

- Screened and selected ~200 leader candidates, standardizing evaluation criteria and coordinating events with the Student Support Office.
- Mentored ~30 leaders to 100% on-time delivery of program reports and deliverables.

PROJECTS

AskData, a Self-Verifying Text-to-SQL Agent 2026
Python • DuckDB • LLM APIs (OpenAI, Anthropic, Ollama)

- Built an LLM agent that answers questions over a database and checks each answer with schema validation, result sanity checks, and an independent LLM judge, labeling every answer trusted, repaired, or abstained instead of guessing.
- Evaluated on a 150-question golden set across five models, three runs each, with exact McNemar significance tests: verification cut confidently-wrong answers by up to 69% (21.8% to 6.7%).
- Found verification quality depends on the judge more than the generator; swapping only the judge model cut a weak generator's error rate five-fold. Study written up as a paper draft.
- Open-source project • live demo at askdata-arthanz.streamlit.app.

BigQuery • Looker Studio • SQL

- Designed a star-schema dimensional model over a large e-commerce dataset (Olist) and built an ETL pipeline into BigQuery following Kimball methodology.
- Delivered an interactive Looker Studio dashboard tracking sales, delivery-performance, and customer-satisfaction KPIs for non-technical stakeholders.
- Documented metric definitions and data lineage to support governance and reproducibility.

Python • PyTorch • Hugging Face Transformers

- Audited the source benchmark (IEEE WCCI/IJCNN 2022) before training and found its “non-advert” Twitter split still full of advertisements; restricted the study to the usable Reddit subset.
- Fine-tuned and benchmarked BERT, RoBERTa, and DistilBERT for binary stress detection; RoBERTa reached 97.2% accuracy and 0.997 ROC-AUC.
- Assessed accuracy-vs-efficiency trade-offs (DistilBERT ~2× faster inference) using macro / weighted F1, per-class metrics, and confusion matrices to guide deployment choices.

Python • scikit-learn • LightGBM • XGBoost

- Built an end-to-end classical ML pipeline (TF-IDF features, SMOTE, hyperparameter tuning) comparing seven models on a severely imbalanced ~60k-sample Indonesian dataset.

PUBLICATIONS & RESEARCH

- Bara, A. N., et al. “Sentiment Analysis of Jakarta Restaurant Reviews Using Classical Machine Learning and TF-IDF.” *IEEE EECCIS 2026* (in press).
- Buananta, S. E. A., Hasani, M. F., Bara, A. N., & Wijaya, H. “User Experience Analysis when Finding Information: Comparative Study of RAG and Traditional Search Engine.” *2024 Int’l Conference on Informatics, Multimedia, Cyber and Information System (ICIMCIS)*, IEEE, 2024, pp. 798–803. doi: 10.1109/ICIMCIS63449.2024.10957499.

CERTIFICATIONS

- Alibaba Cloud Certified Associate (ACA), Cloud Engineer • Alibaba Cloud (valid to Feb 2028).
- Data Analysis Using Python • Cognitive Class (IBM).

AWARDS & ACTIVITIES

- Favorite Presentation Winner, Pekan Ilmiah Mahasiswa Nasional (PIMNAS) ke-36 (2023).
- Grand Finalist & Most Favorite Winner, SOCS Hackathon (2023).
- SASC Mentor Scholarship Awardee (2023) • Finalist, Duta Bahasa BINUS (2023).

SKILLS

Data & BI	SQL, BigQuery, Looker Studio dashboards, dimensional modeling (Kimball / Inmon), star schema, ETL pipelines, KPI reporting, data visualization
Analytics & Statistics	Python (pandas, scikit-learn, PyTorch, TensorFlow), data cleaning & wrangling, hypothesis testing & significance analysis (McNemar, paired tests), A/B-test evaluation, classical ML, NLP
GenAI / LLM	Text-to-SQL agents, LLM-as-judge evaluation pipelines, golden-set evaluation design, prompt engineering
Professional	Stakeholder & sponsor management, data storytelling, public speaking, cross-functional coordination, project management
Tools	Git / GitHub, Google Cloud Platform, Excel, Jupyter, Microsoft Office

LANGUAGES

- Indonesian (native) • English (professional working proficiency, TOEFL 603) • Mandarin (beginner)