Bintang Arigo Kautsar Urumsah

Yogyakarta, Indonesia

arigourumsah.my.id | github.com/arigourumsah | linkedin.com/in/arigour

arigourumsah@gmail.com

Interdisciplinary MSc graduate bridging computer science and international economics, with experience in machine learning, NLP, econometrics, and data analysis using Python, SQL, and Stata. Interested in applying computational and quantitative methods to study technology diffusion and how digitalisation shapes workers, firms, and economies.

EDUCATION

Corvinus University of Budapest - Budapest, Hungary **2022 - 2025**
Master's Degree in International Economy and Business
Specialised in Economic Analysis. Studying under the Stipendium Hungaricum Scholarship - 4.54/5.00 GPA

Gadjah Mada University - Yogyakarta, Indonesia **2015 - 2019**
Bachelor's Degree in Computer Science
Specialised in Data Science and Machine Learning. Graduated with Honours - 3.42/4.00 GPA

RESEARCH EXPERIENCE

Master's Thesis **2024**
Corvinus University of Budapest
Investigating the Multifaceted Factors Influencing Information and Communications Technology (ICT) Growth.

- Combined World Bank and V-Dem data to examine socioeconomic and political correlates of internet adoption; the main model sample contained 2,592 observations across 124 countries.
- Estimated country and year fixed-effects models in Stata with clustered robust standard errors, supplemented by alternative specifications, dynamic models, and regional and period-specific analyses.
- Found that GDP per capita, FDI, and educational inequality were positively associated with internet usage, while political competition was negatively associated, with relationships varying across periods and regions.

Data Analysis for World Economics Course Research Project **2023**
Corvinus University of Budapest
An Empirical Analysis of Factors Impacting Political Stability.

- Co-authored a study of political stability and economic, educational, and institutional factors using country-level panel data covering 202 countries.
- Contributed to fixed-effects panel regression with population weights and robust standard errors, including model comparisons, regional analysis, and robustness checks.
- Found positive associations with GDP per capita, domestic autonomy, and educational equality and negative associations with suffrage and coup attempts, with variation across regions.

Bachelor's Thesis **2019**
Universitas Gadjah Mada
Opinion Mining Using Naive Bayes Algorithm on Online Accommodation Customer Reviews.

- Collected and manually labelled 750 English-language Airbnb reviews into balanced positive, negative, and neutral classes.
- Developed a three-class opinion-mining pipeline in Python, implementing TF–IDF weighting and Naive Bayes classifier from scratch rather than using pre-built machine-learning estimators.
- Improved three-class performance through modified stop-word handling, achieving 72% accuracy, 75% precision, 72% recall, and a 73% F-measure.

SELECTED ACADEMIC OUTPUT

Research Design and Methods for World Economics Course Policy Brief Project **2023**
Corvinus University of Budapest
Impact of ChatGPT in Education: How Can We Balance the Benefits and Risks of ChatGPT?

- Co-authored a policy brief using literature review, comparative policy analysis, case studies, and a 61-respondent questionnaire to assess ChatGPT's educational benefits and risks.
- Developed policy recommendations and co-presented an accompanying research poster as the final course deliverable.

DATA SCIENCE & ML PROJECTS

Laptop Pricing Analytics and Predictive Modelling **2026**

- Developed an ML regression model to estimate laptop prices across seven algorithms, with XGBRegressor achieving an R^2 of 0.897 and RMSE of 225.51 via feature engineering and preprocessing pipelines.

Telecommunication Customer Churn Prediction **2026**

- Developed a customer churn prediction model using supervised machine learning, feature engineering, SMOTE, and hyperparameter tuning, achieving an ROC-AUC of 0.75 and an F1-score of 0.61.

Multi-Class Sentiment Classification of Product Reviews Using Classical NLP Models **2026**

- Built a multi-class sentiment classifier for product reviews using text preprocessing and TF-IDF across five algorithms, with Logistic Regression achieving a peak Macro F1-score of 0.655 via cross-validated tuning.

PROFESSIONAL EXPERIENCE

TellarTech **December 2024 - Present**
Data Scientist and Content Creator **Indonesia (Remote)**

- Lead development of an end-to-end Python and SQL-based content-recommendation system: define data contracts, build data pipelines, design a database, and integrate a backend API for a 460+ game catalogue.
- Build a hybrid recommendation system using game metadata, categories, and TF-IDF features to generate personalised top-3 recommendations for registered users, with mean computation time below 1 ms.
- Co-lead strategy across 3 social-media platforms, delivering 2-4 long-form and 8+ short-form assets monthly while coordinating 2 contributors; grew YouTube to 409K+ views, 5.1K+ watch hours, and 880 subscribers.

Plataran Indonesia **August 2023 - May 2025**
Data Analyst - Food & Beverage Operations **Budapest, Hungary (Part-time)**

- Designed and automated data pipelines and dashboards in Excel and Google Sheets to support daily ops and decision-making for a large team, cutting reporting time by 40% through streamlined workflows.
- Drove monthly financial and operational analyses to identify cost-saving opportunities, contributing to a 15% improvement in net margins while communicating insights to mixed technical and non-technical stakeholders.
- Developed an interactive Google Data Studio dashboard to visualise restaurant sales and employee performance, enabling stakeholders to monitor business performance and make data-driven decisions.

PT. Gosantha Global Indonesia **August 2019 - August 2022**
Software Engineer **Yogyakarta, Indonesia**

- Built and customised Python and SQL-based ETL pipelines on Odoo ERP, automating large-scale workflows for 100+ users and reducing manual reporting time by 30% through robust integration and schema standardisation.
- Led cross-functional integration of POS with 5+ modules, unifying product, customer, and transaction datasets; standardised schemas reduced reconciliation by 25% and enabled accurate sales performance reporting.
- Analysed 10,000+ transactional records to identify and resolve POS integration bottlenecks, optimising data flow and improving system responsiveness by 20%.
- Managed VMware and Ubuntu-based infrastructure supporting ERP data pipelines, ensuring 99.9% system availability and reliable data processing for enterprise reporting.

AWARDS

Stipendium Hungaricum Scholarship - Tempus Public Foundation, fully funded MSc scholarship **2022–2024**
Beasiswa Pendidikan Indonesia (BPI) - Kemendikbudristek, supplementary MSc scholarship **2022–2024**

TECHNICAL SKILLS

Programming: Python, Pandas, NumPy, Scikit-learn, NLTK, SQL (PostgreSQL, MySQL), JavaScript, XML, Git
Machine Learning & Statistical Learning: Supervised learning (classification, regression), predictive modelling, feature engineering, model evaluation, cross-validation, hyperparameter optimisation, ensemble methods, imbalanced learning.
Statistical & Econometric Methods: Statistical analysis, panel data analysis, fixed-effects regression, econometric modelling, statistical inference.
Natural Language Processing: Text classification, sentiment analysis, TF-IDF, text preprocessing, Naive Bayes, Logistic Regression, Linear SVM.
Data Engineering & Data Analysis: ETL pipelines, data integration, data cleaning, data transformation, database management, data visualisation.
Software & Tools: STATA, Tableau, Excel, Google Sheets, Google Data Studio, FastAPI, Flask, VMware

LANGUAGE SKILLS

English and **Indonesian** - Native
Japanese - Conversational
Hungarian - Basic