Email: **gerrydito@apps.ipb.ac.id** Website: **gerrydito.github.io/about**

Gerry Alfa Dito

Lecturer in Statistics and Data Science

EDUCATION

Magister – Statistika
IPB University Sep 2017 — Jan 2020
Thesis: ”Study of Super Learner Development with Cuckoo Search for Increasing Classification Performance”
Supervisor: Dr. Bagus Sartono, Dr. Eng. Annisa, Dr. Anang Kurnia

Sarjana – Matematika
Lampung University Sep 2012 — Apr 2016
Specialization in Statistics
Bachelor’s thesis: ”Neural Network Fuzzy Learning Vector Quantization (FLVQ) to Identify Probability Distributions”
Supervisor: Dr. Warsono and Dr. Dian Kurniasari

Certified Data Analyst
Global ICT Professional Certification Institution Jul 2024 — Jul 2027
View Certification

WORK EXPERIENCE

Lecturer in Statistics and Data Science
School of Data Science, Mathematics and Informatics — Apr 2020 – Present
IPB University

Contributing Write
Info Komputer — Oct 2019 – Mar 2020

Research Assistant
Department of Statistics — Oct 2019 – Mar 2020
IPB University

Statistical Consultant
Dokter Data — Dec 2017 – Oct 2019
Lembaga Ilmu Pengetahuan Indonesia(LIPI) — Sep 2018 – Oct 2018
Detective Data — Sep 2016 – Aug 2017

Teaching Assistant
Department of Mathematics — Sep 2013 – Jan 2016
Lampung University

Data Analyst
Representative Office of Bank Indonesia Lampung — Jan 2015 – Feb 2015

RESEARCH INTEREST

My research interests lie at the intersection of statistical theory, machine learning, and applied data science, focusing on both theoretical advancements and practical applications. Specifically, my work addresses the following areas:

1. **Bayesian Methods**: Development of scalable Bayesian inference techniques, with a focus on Markov Chain Monte Carlo (MCMC) methods and Variational Inference.
2. **Interpretable Machine Learning**: Developing machine learning models that are inherently understandable by humans. This involves designing algorithms where the decision-making process is transparent and directly comprehensible, such as sparse linear models, decision trees with a limited number of nodes, and rule-based models. My research in this area emphasizes the creation of models that maintain high performance while ensuring that the underlying mechanisms and predictions are intuitive and easily interpretable. This is particularly important in high-stakes applications such as healthcare and finance, where trust and accountability are crucial.

3. **Explainable Machine Learning**: Developing methods to enhance model transparency and interpretability in complex machine learning models, particularly deep learning and ensemble methods. This includes both post-hoc interpretability techniques and the development of inherently interpretable models.
4. **Applications in Biostatistics, Economics, and Social Sciences**: Applying advanced statistical and machine learning techniques to real-world problems in health, economics, and social science.

SELECTED PUBLICATIONS

Khikmah, K. N., Sartono, B., Susetyo, B., & Dito, G. A. (2024). Performance Comparative Study of Machine Learning Classification Algorithms for Food Insecurity Experience by Households in West Java. *Jurnal Online Informatika*, *9*(1), 128–137. https://doi.org/10.15575/join.v9i1.1012

Sartono, B., Trisnawati, D. A., Dito, G. A., Hadi, A. F., & Ramadhani, E. (2024). Recursive Feature Elimination with Cross-Validation in Support Vector Machine for Recognizing Food Insecurity Cases in West Java. In *The 18th IMT-GT International Conference on Mathematics, Statistics and their Applications* (pp. 75–79). Sciendo. https://doi.org/10.2478/9788367405713-014

Yanti, Y., Rahardiantoro, S., & Dito, G. A. (2023). Spatio-temporal Clustering Analysis of Dengue Hemorrhagic Fever Cases in West Java 2016 – 2021: Analisis Penggerombolan Spasio-temporal Kasus DBD di Jawa Barat Tahun 2016 – 2021. *Indonesian Journal of Statistics and Its Applications*, *7*(1), 56–63. https://doi.org/10.29244/ijsa.v7i1p56-63

Dito, G. A., Sartono, B., & Annisa, A. (2020). Super Learner for Predicting Stock Market Trends: A Case Study of Jakarta Islamic Index Stock Exchange. *Proceedings of the Proceedings of the 1st International Conference on Statistics and Analytics, ICSA 2019, 2-3 August 2019, Bogor, Indonesia*. https://doi.org/10.4108/eai.2-8-2019.2290523

Sartono, B., Bodro, D. K., & Dito, G. A. (2020). *Teknik eksplorasi data yang harus dikuasai data scientist*. IPB Press.

Dito, G. A., Safitri, A., Afendi, F. M., Anisa, R., Salim, A., & Sartono, B. (2019). Graphical user interface (GUI) for the least absolute shrinkage and selection operator (LASSO) regression. *IOP Conference Series: Earth and Environmental Science*, *299*(1), 012031. https://doi.org/10.1088/1755-1315/299/1/012031

Warsono, W., Dito, G. A., Kurniasari, D., & Usman, M. (2016). Neural Network Fuzzy Learning Vector Quantization (FLVQ) to Identify Probability Distributions. *International Journal of Computer Science and Network Security*, *16*(10), 16–19.

SOFTWARE AND TECHNICAL SKILLS

- **Programming Languages**:
  - **R**: Extensive experience in data analysis, statistical modeling, and visualization using R. Proficient with libraries such as `tidyverse` ecosystems.
  - **Python**: Skilled in Python for data science and machine learning, using libraries such as `pandas`, `NumPy`, `SciPy`, and machine learning tools like `Scikit-Learn`, `XGBoost`, and `matplotlib`.
- **Statistical Software**:
  - **SAS**: Proficient in data manipulation, statistical analysis, and predictive modeling using SAS. Experienced with SAS procedures for regression, time series analysis, and generalized linear models.
- **Machine Learning Frameworks**:
  - **Scikit-Learn**: Expertise in implementing machine learning algorithms, including classification, regression, clustering, and model evaluation techniques.
  - **Tidymodels**: Skilled in using the Tidymodels framework for modeling in R, including feature engineering, cross-validation, and building machine learning pipelines.
- **Deep Learning Frameworks**:
  - **TensorFlow**: Experienced in building, training, and deploying deep learning models for tasks such as image recognition, natural language processing, and time-series forecasting.
  - **PyTorch**: Proficient in developing and fine-tuning deep learning models using PyTorch, with applications in neural networks, convolutional networks, and reinforcement learning.
- **Bayesian Modeling**:
  - **Stan**: Proficient in Bayesian inference and probabilistic programming using Stan, with experience in hierarchical modeling, MCMC methods, and advanced sampling techniques.

CONSULTING AND INDUSTRY COLLABORATION

Workshop Instructor – Pendidikan In House Training (IHT) Kaidah Penyusunan Riset, *BRI Corporate University, PT. Bank Rakyat Indonesia (Persero) Tbk* – 23 Aug 2024

- Developed custom training materials focused on **data exploration**, **statistical modeling**, and **interpretation of results** for business banking applications using Jamovi.
- Delivered hands-on workshops on **data exploration**, **statistical modeling**, and **interpretation of results** for business banking applications using Jamovi.

Data Science Consultant – *PT. Bumitama Gunajaya Argo* – Aug 2023 to Dec 2023

- Developed forecasting models for FFB (Fresh Fruit Bunch) Production in Oil Palm Plantations using XGBoost algorithms.
- Utilized R to create data pipelines, perform feature engineering, and deploy XGBoost models algorithms to forecast FFB (Fresh Fruit Bunch) Production in Oil Palm Plantations.

Workshop Instructor – Bimbingan teknis pengolahan dan analisis data spasial untuk wilayah provinsi atau kabupaten/kota, *Jurusan Kepabeanan dan Cukai, Politeknik Keuangan Negara STAN* – 17 Mar 2021

- Developed custom training materials focused on **spatial data exploration** and **spatial regression modeling** using R.
- Delivered hands-on workshops on **spatial data exploration** and **spatial regression modeling** using R.

Data Science Consultant – *Ministry of Villages, Development of Disadvantaged Regions, and Transmigration* – Mar 2022 to Aug 2022

- Developed models for identification Identifying factors that affect household poverty status.
- Developed models for effective policy interventions for extreme poverty alleviation

Data Science Consultant – *Otoritas Jasa Keuangan (OJK)* – Aug 2020 to Jan 2021

- Developed a machine learning model for monitoring crisis management protocol indicators in the stock market.
- Utilized R to create machine learning pipelines to monitor crisis management protocol indicators in the stock market.

PROFESSIONAL DEVELOPMENT

Certification

1. **How to Write a Successful Research Paper**
   Udemy (2024)
   Completed courses on UNDERSTAND the logic and structure of a research paper, IDENTIFY the qualities that make a research paper effective.
2. **The Data Science Course: Complete Data Science Bootcamp**
   Udemy (2022)
   Completed courses on data science field and the type of analysis carried out, Mathematics , Statistics, Python, Data Visualization, Machine Learning, Deep Learning
   View Certification
3. **Deep Learning A-Z: Hands-On Artificial Neural Networks**
   Udemy (2021)
   Completed courses on neural networks, CNNs, RNNs, Self-Organizing Maps, Autoencoder, and practical deep learning projects using TensorFlow and PyTorch.
   View Certification