

ARYO PRADIPTA GEMA

<https://aryopg.github.io/> • +44 7476 68 68 38 • [aryopg\[at\]gmail\[dot\]com](mailto:aryopg[at]gmail[dot]com)

CURRENT RESEARCH INTERESTS

Primary : Biomedical NLP, Language Modelling, Neuro-symbolic AI
Secondary : Computational Biology, Argumentation Mining, Generative Adversarial Network

PUBLICATIONS

- Gema, A., Daines, L., Minervini, P., Alex, B. (2023). **Parameter-Efficient Fine-Tuning LLaMA for the Clinical Domain**. arXiv preprint arXiv:2307.03042.
- Gema, A., Grabarczyk, D., De Wulf, W., Borole, P., Alfaro, J., Minervini, P., Vergari, A., & Rajan, A. (2023). **Knowledge Graph Embeddings in the Biomedical Domain: Are They Useful? A Look at Link Prediction, Rule Learning, and Downstream Polypharmacy Tasks**. arXiv preprint arXiv:2305.19979.
- Gema, A., Kobiela, M., Fraise, A., Rajan, A., Oyarzun, D., & Alfaro, J. (2023). **Vaxformer: Antigenicity-controlled Transformer for Vaccine Design Against SARS-CoV-2**. arXiv preprint arXiv:2305.11194.
- Halimawan, N., Suhartono, D., Gema, A., & Yunanda, R. (2022). **BERT and ULMFiT Ensemble for Personality Prediction from Indonesian Social Media Text**. In 2022 International Symposium on Information Technology and Digital Innovation (ISITDI) (pp. 156–161).
- William, D., Achmad, S., Suhartono, D., & Gema, A. (2022). **Leveraging BERT with Extractive Summarization for Depression Detection on Social Media**. In 2022 International Seminar on Intelligent Technology and Its Applications (ISITIA) (pp. 63–68).
- Saputra, F., Namyu, U., Suhartono, D., & Gema, A. (2021). **Automatic Piano Sheet Music Transcription with Machine Learning**. Journal of Computer Science, 17(3), 178–187.
- Suhartono, D., Gema, A., Winton, S., David, T., Fanany, M., & Arymurthy, A. (2020). **Sequence-To-Sequence Learning For Motion-Aware Claim Generation**. International Journal of Computing, 19(4), 620–628.
- Suhartono, D., Gema, A., Winton, S., David, T., Fanany, M., & Arymurthy, A. (2020). **Argument annotation and analysis using deep learning with attention mechanism in Bahasa Indonesia**. Journal of Big Data, 7, 1–18.
- Suhartono, D., Gema, A., Winton, S., David, T., Fanany, M., & Arymurthy, A. (2019). **Attention-based argumentation mining**. International Journal of Computational Vision and Robotics, 9(5), 414–437.
- Anderson, R., Gema, A., Isa, S., & others (2018). **Facial attractiveness classification using deep learning**. In 2018 Indonesian Association for Pattern Recognition International Conference (INAPR) (pp. 34–38).
- Suhartono, D., Gema, A., Winton, S., David, T., Fanany, M., & Arymurthy, A. (2017). **Hierarchical attention network with XGBoost for recognizing insufficiently supported argument**. In Multi-disciplinary Trends in Artificial Intelligence: 11th International Workshop, MIWAI 2017, Gadong, Brunei, November 20-22, 2017, Proceedings 11 (pp. 174–188).
- Gema, A., Winton, S., David, T., Suhartono, D., Shodiq, M., & Gazali, W. (2017). **It takes two to tango: modification of siamese long short term memory network with attention mechanism in recognizing argumentative relations in persuasive essays**. Procedia computer science, 116, 449–459.

EDUCATION

UNIVERSITY OF EDINBURGH

Edinburgh, Scotland

- PhD in Biomedical AI (Incoming) September 2023 - Present
 - A part of the fully-funded four-year studentship, CDT in Biomedical AI.
 - A member of the Clinical NLP group (co-led by Dr Beatrice Alex and Dr Honghan Wu).
 - Supervised by Dr Beatrice Alex, Dr Pasquale Minervini, and Dr Luke Daines
- MSc by Research in Biomedical AI September 2022 - August 2023
 - Thesis: Parameter-efficient Fine-tuning of Large Language Models for the Clinical Domain.
 - A part of the fully-funded four-year studentship, CDT in Biomedical AI.
 - A member of the Clinical NLP group (co-led by Dr Beatrice Alex and Dr Honghan Wu).
 - Supervised by Dr Beatrice Alex, Dr Pasquale Minervini, and Dr Luke Daines

BINA NUSANTARA UNIVERSITY

Jakarta, Indonesia

- Bachelor of Science in Computer Science September 2014 - August 2018
 - Thesis: Sequence-To-Sequence Learning For Motion-Aware Claim Generation
 - A-Accredited university for Computer Science Undergraduate Program in Indonesia.
 - Concentration: Artificial Intelligence (**current GPA: 3.87/4.00**).

PROFESSIONAL EXPERIENCES

EPIGENE LABS (AI-driven Oncology startup)

Paris, France

- Senior Data Scientist October 2021 - August 2022
 - Develop an internal **source-agnostic big data management end-to-end pipeline**.
 - Maintain the quality of **more than 20 Machine Learning models** in production.
 - Mentor Junior data scientist(s).
 - Work with the Head of Data to align Data Science R&D and product timeline.
 - Initiated a regular **inter-departmental research workshop**.
- Data Scientist October 2019 - September 2021
 - Research and Implementation of **Language-Agnostic, Format-Agnostic, and Cancer-Agnostic Data Harmonisation Pipeline** using hybrid **Machine Learning** techniques.
 - Maintain the quality of **more than 20 Machine Learning models** in production.
 - Supervise data collection, data labelling, and the formulation of ontology.
 - Supervise Data Science internship projects on Tailored Performance Metrics and Continuous Learning.
 - Contribute to **an open-sourced Computational Biology project** (<https://github.com/epigenelabs/pyComBat>).
 - Directly worked under the C-level to align data science research, DevOps, and product timeline before the arrival of the Head of Data.

RIMINDER (Deep Learning based HR startup)

Paris, France

- Staff AI Research Scientist February 2019 - October 2019
 - **Supervised day-to-day operations** of the R&D team.
 - Directly worked with the C-level to align R&D and product timelines.
 - Improved in-house NER model by 20% F1 score, block tagger, and other predictive models by 5%.
 - **Double promotion after three months** of the trial period after successfully **improving in-house recommendation and NLU algorithms**.
- AI Research Scientist October 2018 - January 2019

- Contributed to the research team to further improve profile-to-job **matching/recommendation** using deep learning algorithms.
- Introduced novel approaches for profile-to-job matching, and improved the model by a 10% ROC AUC score.
- **Held ownership** of multiple of Riminder's most significant projects.
- Primarily used Python and its libraries (Tensorflow, PyTorch, Keras, NLTK, Pandas, Scikit-learn, Numpy).

JULO (AI-driven fin-tech startup)

Jakarta, Indonesia

- **Data Scientist** September 2017 - July 2018
 - Developed multiple machine learning solutions, including Automatic individual **credit performance grading, deep learning-based supervised and unsupervised text processing, and recognition**.
 - Worked collaboratively with the Head of Analytics.
 - Mentored junior data scientists and interns.
 - Primarily used Python and its libraries (Tensorflow, Pytorch, H2O, Keras, NLTK, Pandas, Scikit-learn, Numpy).

BINA NUSANTARA UNIVERSITY (private university)

Jakarta, Indonesia

- **Junior Deep Learning Research Scientist** March 2017 - July 2018
 - Researched cutting-edge **Deep Learning** and other **Machine Learning** techniques in the **Argumentation Mining** domain.
 - Published three papers in two different international conferences and a journal (**ICCCSCI 2017, MIWAI-SoDeLIA 2017, IJCVR 2018**) as of April 2018. Presented one of the papers in a national workshop (**INACL 2017**).
 - Best Presenter of ICCSCI 2017.
 - Appointed as the **lead researcher** for the team and managed fellow junior researchers' workload.
 - Primarily used Python and its libraries (Tensorflow, PyTorch, Keras, NLTK, Pandas, Scikit-learn, Numpy).

RE/MAX 2000 (real-estate company)

California, USA

- **Freelance Web Developer** July 2017 - September 2017
 - Developed 10 core modules for internal **Customer Relationship Management (CRM)** application.
 - Developed 8 core modules for **Image and Descriptive Text Verification** applications.
 - Developed 5 WordPress-based websites.

VOLANTECH (B2B multi-platform applications start-up)

Jakarta, Indonesia

- **Senior Software Engineer** February 2016 - July 2017
 - Managed and developed multiple high-value projects at the same time.
 - Worked closely with the C-level.
 - Research on **Deep Learning** techniques in conjunction with **Internet of Things** systems.
 - Formulated **competitive programming quizzes** for interview purposes and became the interviewer as well.
 - Primarily used Laravel (PHP), Meteor.js, MEAN stack, Google Firebase, and Python.

AWARDS

- **Academic**
 - **Binusian Award 2018** (*awarded to the most achieving graduate*)
 - **“Penelitian Produk Terapan (Applied Research Grant) Binus” 2017.** (*awarded to research initiatives by Binus University*)
 - **Scholarship Mentor Program 2015** (*awarded to students that successfully mentor their peers*)
 - **Student of Excellence Honour 2015** (*awarded to students with international and national achievements*)
 - **Student of Excellence Honour 2014** (*awarded to students with international and national achievements*)

- **Non-academic**

- Quarterfinalist in English as Foreign Language Category of United Asian Debating Championship 2015, Indonesia.
- Champion in Novice Category of Indonesian Varsity English Debate 2015, Indonesia.
- Semifinalist and 8th Best Speaker of Universitas Padjadjaran Debating Championship 2015.
- 1st Runner Up and 3rd Best Speaker of Newbies Debating Championship 2014.

INVITED TALKS

- | | |
|--|-----------------------------------|
| ● Knowledge Graph Embeddings in the Biomedical Domain | Binus University, May 2023 |
| ● Knowledge-Augmented Clinical Language Models | Binus University, May 2023 |
| ● ChatGPT: A Biomedical AI Perspective | University of Edinburgh, May 2023 |
| ● Knowledge-Augmented Clinical Language Models | Binus University, March 2023 |
| ● Clinical Language Models | Binus University, February 2023 |
| ● Building a Knowledge Base from Biomedical Literature | Binus University, July 2022 |
| ● Augmenting Intelligence to Drug Discovery | Binus University, December 2021 |
| ● Augmenting Intelligence to Drug Discovery | Binus University, April 2021 |
| ● Introduction to Artificial Intelligence | Binus University, October 2020 |

SKILLS AND INTERESTS

- **Languages:** Indonesian – Native, English – Fluent, French – Basic
- **Skills:** Python, Golang, C++, Tensorflow, Pytorch, Scikit-Learn, spaCy, NLTK, Numpy, Scipy, Pandas, H2O, MySQL, PostgreSQL, MongoDB, Azure Cosmos DB
- **Other Interests:** Chess, Parliamentary Debate, Sports, Cooking