DYAH ADILA

Research Interests *Foundation Models, Robustness, Learning with limited data*

Contact Information *E-mail:* adila@wisc.edu
Website: dyahadila.github.io

Education

Ph.D. in Computer Science University of Wisconsin - Madison **Sep 2021 - Present**
- Advisor: Fred Sala

M.S. in Computer Science University of Minnesota - Twin Cities **Sep 2019 - May 2021**
- Advisor: Ju Sun

B.Eng. Nanyang Technological University (NTU), Singapore **Aug 2013 - Jun 2017**

Conference & Journal Publications

Dyah Adila, Zhang. S, Han. B, Wang. Y, *Discovering Bias in Latent Space: An Unsupervised Debiasing Approach*, in International Conference on Machine Learning (ICML), 2024. [Paper] [Code]

Dyah Adila*, C. Shin*, L. Cai, F.Sala, *Zero-Shot Robustification of Zero-Shot Models*, in International Conference on Learning Representations (ICLR) 2024. NeurIPS 2023 R0-FoMo Workshop. **Oral presentation (best paper honorable mention)**. [Paper] [Code] [Blog]

N. Roberts*, X. Li*, **Dyah Adila**, S. Cromp, B. Huang, J. Zhao, F. Sala, *Geometry-Aware Adaptation for Pretrained Models*, in Neural Information Processing Systems (NeurIPS), 2023. [Paper]

C. Shin, S. Cromp, **Dyah Adila**, S., F. Sala, *Mitigating Source Bias for Fairer Weak Supervision*, in Neural Information Processing Systems (NeurIPS), 2023. [Paper]

M. Chen*, D. Fu*, **Dyah Adila**, M. Zhang, F. Sala, K. Fatahalian, C. Ré, *Shoring Up the Foundations: Fusing Model Embeddings and Weak Supervision*, in Uncertainty in Artificial Intelligence (UAI), 2022. **Oral presentation (best student paper runner-up)**. [Paper] [Code]

N. Roberts*, X. Li*, B. Huang, **Dyah Adila**, S. Schoenberg, C. Liu, L. Pick, H. Ma, A. Albarghouthi, F. Sala, *AutoWS-Bench-101: Benchmarking Automated Weak Supervision with 100 Labels*, in Neural Information Processing Systems (NeurIPS), 2022. [Paper]

Workshop Publications

Dyah Adila, C. Shin, Y. Zhang, F. Sala, *Can Language Models Safeguard Themselves, Instantly and For Free?*, in ICML 2024 Next Generation of AI Safety Workshop. [Paper]

Dyah Adila, S. Cromp, S. Mo, F. Sala, *Causal Omnivore: Fusing Noisy Estimates of Spurious Correlations*, in ICML 2022 Workshop on Spurious Correlations, Invariance, and Stability. [Paper]

Dyah Adila, and Dongyeop Kang, *Understanding Out-of-distribution: A Perspective of Data Dynamics*, in NeurIPS 2021 Workshop: I Can't Believe It's Not Better!, PMLR 2022. [Paper]

Awards

Qualcomm Innovation Fellowship Finalist 2024
Best paper award honorable mention at NeurIPS R0-FoMo 2023
ICCV DataComp winning team in small scale filtering track 2023
UAI Best Student Paper Runner-up 2022

Experience

University of Wisconsin-Madison, USA **Aug 2021 - Present**
Research Assistant
- Ph.D. research in Machine Learning advised by Fred Sala.

Amazon Web Services AI, USA **Fall 2023, Summer 2024**
Applied Scientist Intern
- Designed method to denoise retrieval augmented generation (RAG) knowledge base (Summer'24).
- Designed attention steering method to mitigate cognitive and social bias in foundation models, results published in ICML 2024. (Fall'23)

University of Minnesota, USA **May 2020 - May 2021**
Research Assistant
- Developed a rapid diagnostic model for COVID-19 in chest X-rays and built GAN for data augmentation to tackle class imbalance.
- Press coverage: [Link 1] [Link 2] [Link 3]

Traveloka, Indonesia **Jul 2017 - Jul 2019**
Software Engineer
- Led the development of company-wide React Native user interface library (runs on Android and iOS), which speed up development time by 2x.
- Built Traveloka's customer-facing and business-facing mobile applications.

JPMorgan Chase, Singapore **May 2016 - Jul 2016**
Software Engineer Intern
- Built a real-time log monitoring tool to keep track of daily transactions using Java.
- Built an API for multiple currency transfer application.

Seagate Technology, Singapore **Jan 2016 - May 2016**
Research Engineer Intern
- Built a continuous integration framework to automate Seagate's build and test pipelines.

Technical Skills
- **Competent:** Python, PyTorch, TensorFlow, Java, Unix, Google Cloud Platform, JavaScript
- **Familiar:** SQL, AWS, Apache Spark