

Muhammad Hilmi Asyrofi

✉ mhilmiasyrofi@gmail.com | ☎ (+65) 85472451

📄 github.com/mhilmiasyrofi | [in linkedin.com/in/mhilmiasyrofi/](https://www.linkedin.com/in/mhilmiasyrofi/) | [Portfolio](#) | [Google Scholar](#)

🗨 a Kaggle Data Science Expert with 4 years of experience in delivering high-impact works in the IT industry

EXPERIENCES

Machine Learning Engineer

Singapore

PropertyGuru - Data Science Center of Excellence

Jun 2022 - present

- Worked closely with PM and Ops team to understand business problems and translated them into profitable solutions. Collaborated with cross-functional teams to **productionize data science projects as sellable systems to external clients**, e.g. Automated Valuation Model for banks and real estate consultants.
- Developed an Address Matching system to synchronize a physical location with address records in the data warehouse. Leveraged advance natural language processing techniques to **aggregate data from multiple sources to provide rich information from an address**, e.g. property information, property valuation, and property risk assessment. Improved the matching rate by more than 70% compared to the legacy SQL-based string matching.
- **Led cross-team collaboration with data engineers and software engineers** to solve inaccurate address listings in Malaysia. Fixed 150k inaccurate listing data, resulting in more accurate location-based listing recommendations and an estimated six-month revenue increase of over USD100k.
- **Led an initiative to incorporate Trustworthy metrics** in the machine learning systems. Detected and performed root cause analysis on unacceptable fluctuating predictions in a highly critical product. As a result, the volatile predictions were reduced by more than 90%.

AI Research Engineer

Singapore

School of Computing and Information System, Singapore Management University

Nov 2019 - May 2022

- Conducted research on **Trustworthy Artificial Intelligence** supervised by ACM Distinguished Scientist [Prof David Lo](#)
- **Published 7 papers** in international conferences and journals, i.e., IEEE TSE (Q1, CCF Rank A, Impact Factor: 9.32), ACM ESEC/FSE (CORE A*), IEEE ICSME (CORE A), and IEEE SANER (CORE A)
- **Maintained machine learning projects** using Dockerized environments running on Nvidia DGX AI Supercomputer

Co-Founder & CTO

Surabaya, Indonesia

Alinamed - [Pitch Deck](#), [Playstore](#)

Jan 2019 - Jul 2020

- **Managed the end-to-end software development lifecycle** by collaborating across members, including functional requirements analysis, design specification, technical code review, project planning, and implementation of applications
- Managed the selection of programming languages, frameworks, cloud, web servers, databases, and mobile
- Won prestigious startup competitions, e.g., 1st Winner at WMM by Bank Mandiri, 1st Winner at Startup Weekend Surabaya, Winner at Tech-to-Impact by UK-ID Tech Hub. Secured IDR 750 million investment from Telkom Indonesia

Machine Learning Engineer Intern

Jakarta, Indonesia

Bukalapak

May 2018 - Aug 2018

- Developed object detection as a part of the search-by-image feature. Developed character recognition to detect seller's information (e.g., phone number and Instagram username) in the e-commerce product image. Localized text lines using CTPN (Connectionist Text Proposal Network) and combined bounding boxes of words within the same context.

HONORS & AWARDS

- 2022 **Bronze Medal | Top 7% (170 from 2,654) | Evaluating Written Essays | NLP | [Kaggle Leaderboard](#)**
The challenge was to assess the language proficiency of students by analyzing their written essays. I leveraged DeBERTa language models to extract the feature embeddings from the essays. Then, I trained neural network layers to predict the proficiency scores from extracted features.
- 2021 **Silver Medal | Top 5% (101 from 2,426) | Shopee Product Matching | NLP + CV | [Kaggle Leaderboard](#)**
The task was predicting similar e-commerce products by using product images and titles. I utilized EfficientNet and NFNet to extract feature embeddings from images and utilized BERT and TFIDF to extract feature embeddings from texts. I calculated the cosine distance between products and employed k-NN to find similar products
- 2021 **Top 3% (28 from 1,034) | Information Extraction - Shopee Data Science Challenge | [Kaggle Leaderboard](#)**
The challenge was extracting Point of Interest (POI) names and street names from unformatted Indonesia addresses collected by Shopee. I framed the challenge as NER (Named Entity Recognition), where POI and street are the extracted entities. Then, I fine-tuned the IndoNLU model on the customized dataset.
- 2019 **Gojek Upscale 02 for Data Science Enthusiast | [Certificate](#)**
A complete two-day experience from the Gojek data science team. It covered machine learning, data modeling, and Gojek data science best-practices. I was one of 20 chosen participants out of 1,000 applicants
- 2011 **Bronze Medal | Mathematics Competition at 2nd National Science Camp | [Certificate](#)**
Awarded by Ministry of Education - Indonesia

TECHNICAL SKILLS

Programming	Python, C++, Java, PHP, JavaScript, HTML, React, Golang, C, C#
Cloud	AWS Amazon Web Services, GCP Google Cloud Platform
ML Library	Scikit-Learn, PyTorch, TensorFlow, OpenCV
Database	MySQL, BigQuery, PostgreSQL, MongoDB

EDUCATION

Institut Teknologi Bandung (ITB) Bandung, Indonesia
Bachelor of Computer Science 2015 - 2019

- **Cum Laude** with GPA **3.68** / 4.00 ([transcript](#))
- [I was honored as an inspiring graduate](#) because of my outstanding academic and non-academic achievements

PUBLICATIONS

Muhammad Hilmi Asyrofi, Yang Zhou, Imam Nur Bani Yusuf, Hong Jin Kang, Ferdian Thung, and David Lo. “**BiasFinder: Metamorphic Test Generation to Uncover Bias for Sentiment Analysis Systems**”. IEEE Transaction on Software Engineering (Q1, CCF Rank A, Impact Factor 6.11). [PDF](#) [GitHub](#) [DOI](#)

Zhou Yang, Jieke Shi, [Muhammad Hilmi Asyrofi](#), and David Lo. “**Revisiting Neuron Coverage Metrics and Quality of Deep Neural Networks**”. *28th IEEE International Conference on Software Analysis, Evolution and Reengineering (SANER)*. Honolulu, Hawaii. March 2022. [PDF](#) [GitHub](#) [DOI](#)

[Muhammad Hilmi Asyrofi](#), Zhou Yang, Jieke Shi, Chu Wei Quan, and David Lo. “**Can Differential Testing Improve Automatic Speech Recognition Systems?**”. 37th IEEE International Conference on Software Maintenance and Evolution (ICSME). Luxembourg. September 2021. [PDF](#) [GitHub](#) [DOI](#)

Zhou Yang, Harshit Jain, Jieke Shi, [Muhammad Hilmi Asyrofi](#), and David Lo. “**BiasHeal: On-the-Fly Black-Box Healing of Bias in Sentiment Analysis Systems**”. 37th IEEE International Conference on Software Maintenance and Evolution (ICSME). Luxembourg. September 2021. [PDF](#) [GitHub](#) [DOI](#)

[Muhammad Hilmi Asyrofi](#), Yang Zhou, and David Lo. “**CrossASR++: A Modular Differential Testing Framework for Automatic Speech Recognition**”. *29th ACM Joint European Software Engineering Conference and Symposium on the Foundations of Software Engineering (ESEC/FSE '21)*. Athens, Greece. August 2021. [PDF](#) [GitHub](#) [Video](#) [DOI](#)

Zhou Yang, [Muhammad Hilmi Asyrofi](#), and David Lo. “**BiasRV: Uncovering Biased Sentiment Predictions at Runtime**”. *29th ACM Joint European Software Engineering Conference and Symposium on the Foundations of Software Engineering (ESEC/FSE '21)*. Athens, Greece. August 2021. [PDF](#) [GitHub](#) [Video](#) [DOI](#)

[Muhammad Hilmi Asyrofi](#), Ferdian Thung, David Lo, and Lingxiao Jiang. “**CrossASR: Efficient Differential Testing of Automatic Speech Recognition via Text-To-Speech**”. *36th IEEE International Conference on Software Maintenance and Evolution (ICSME)*. Adelaide, Australia. October 2020. [PDF](#) [GitHub](#) [Video](#) [DOI](#)

[Muhammad Hilmi Asyrofi](#), Ferdian Thung, David Lo, and Lingxiao Jiang. “**AUSearch: Accurate API Usage Search in GitHub Repositories with Type Resolution**”. *27th IEEE International Conference on Software Analysis, Evolution and Reengineering (SANER)*. Ontario, Canada. February 2020. [PDF](#) [GitHub](#) [Video](#) [DOI](#)

MACHINE LEARNING CERTIFICATIONS

Deeplearning.AI	Neural Network and Deep Learning Credential ID: D8GK6SD6Y5TG
Deeplearning.AI	Improving Deep Neural Networks: Hyperparameter Tuning, Regularization and Optimization Credential ID: SGC98U3G8K6P

INTERPERSONAL SKILLS

- 2018 **Granted Scholarship** from Rumah Amal Salman ITB. This scholarship focused on **capacity building** to create future leaders. I was awarded as **top 10 awardees** of this scholarship
- 2017 Selected participants to take **leadership training** from Siam Cement Group (SCG). Designed by SCG Learning, the program aims to give participants **fundamental leadership concepts** and provide them opportunities to **practice the interpersonal skill** with other young leaders

I declare the details stated to be true and complete.
(This resume is updated on May 9th, 2023)