

## Muhammad Hilmi Asyrofi

✉ [mhilmiasyrofi@gmail.com](mailto:mhilmiasyrofi@gmail.com) | ☎ (+65) 85472451

📄 [github.com/mhilmiasyrofi](https://github.com/mhilmiasyrofi) | 🔗 [linkedin.com/in/mhilmiasyrofi/](https://www.linkedin.com/in/mhilmiasyrofi/) | 📁 [personal projects](#)

🗨 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

- Developed an Automatic Valuation Model that analyzes historical sales transactions, property characteristics, and price trends to estimate the current market value for a specific property. Contributed to the company's revenue by **productionizing the solution as a sellable system to external clients**, including 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: 6.11), ACM ESEC/FSE (CORE A\*), IEEE ICSME (CORE A), and IEEE SANER (CORE A)
- **Maintained** PyTorch and TensorFlow **machine learning projects** on Docker running on Nvidia DGX AI Supercomputer

### Co-Founder & CTO

Surabaya, Indonesia

Alinamed - [Company Profile](#), [Playstore](#), [AppStore](#)

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., [1<sup>st</sup> Winner at WMM by Bank Mandiri](#), 1<sup>st</sup> 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 **Southeast Asia Machine Learning School (SEA MLS)**  
A five-day ML training event held by [Southeast Asia's AI experts](#) working at DeepMind UK and Google Brain USA. I was one of 200 chosen participants out of 1200 applicants across Southeast Asia and India
- 2019 **Gojek Upscale 02 for Indonesia's 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 2<sup>nd</sup> National Science Camp | [Certificate](#)**  
Awarded by Ministry of Education - Indonesia

## TECHNICAL SKILLS

---

<b>Programming</b>	Python, C++, Java, PHP, Javascript, HTML, React, Golang, C, C#
<b>ML Library</b>	PyTorch, Tensorflow, Keras, OpenCV, Scikit-Learn
<b>Cloud</b>	AWS, GCP, Microsoft Azure
<b>Database</b>	MySQL, PostgreSQL, MongoDB

## EDUCATION

---

<b>Institut Teknologi Bandung (ITB)</b>	Bandung, Indonesia
Bachelor of Computer Science	2015 - 2019
<ul style="list-style-type: none"><li>• <b>Cum Laude</b> with GPA 3.68 / 4.00 (<a href="#">transcript</a>)</li><li>• <a href="#">I was honored as an inspiring graduate</a> because of my outstanding academic and non-academic achievements</li></ul>	

## MACHINE LEARNING CERTIFICATIONS

---

<b>Deeplearning.AI</b>	Neural Network and Deep Learning Credential ID: <a href="#">D8GK6SD6Y5TG</a>
<b>Deeplearning.AI</b>	Improving Deep Neural Networks: Hyperparameter Tuning, Regularization and Optimization Credential ID: <a href="#">SGC98U3G8K6P</a>

## 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](#)

## 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 April 2<sup>nd</sup>, 2023)