Rezka Leonandya

Senior ML Engineer (Data Science) with 8+ years of experience spanning both industry and academia.
Specializes in the end-to-end lifecycle of data science.

Personal Website | Github | LinkedIn | Leetcode

INDUSTRY EXPERIENCE

Gojek | The biggest ride-hailing & superapp in Indonesia. $5B USD Valuation. Jakarta, Indonesia
Senior ML Engineer - Routing & Maps Jun 2022 – Present

- Co-led initiative to eliminate dependency on Google Maps, resulting in > 1 million USD in savings annually, by significantly enhancing the internal route quality to reach near parity—within just 1-2% of Google Maps. This was accomplished by training a tree-based gradient boosting model using LightGBM library to predict better time estimate of a given road segments.
- Built and deployed a production-grade internal ETA prediction model leveraging real-time streaming live traffic features, achieving a 50% reduction in ETA error metrics.
- Designed and deployed a production-grade automatic maps error detector to assist human agents in keeping internal maps updated to real-world conditions. This system automated 45% of incoming driver tickets while maintaining over 70% precision by leveraging raw GPS ping data, the map-matching algorithm, and crowd-sourced driver feedback.
- Managed and optimized approximately 200 flyte data production pipelines, both batch and real-time (streaming). The combined pipeline processes over 1 million rows of data daily.
- Designed and ran rigorous experimentation (e.g. AB & AA Test, switchback, and diff-and-diff) to ensure ML model effectiveness and alignment with key business metrics.

HappyFresh | Quick commerce startup. Raised $97M USD in Series D. Jakarta, Indonesia
Data Scientist - Search & Ranking Feb 2021 – Jun 2022

- Improved conversion rate metrics by approximately 2% by enhancing the internal next-purchase ranking models (tree-based model) which predicts better order for items that the users preferred.
- Reduced null search result metrics by around 1% by enhancing the internal synonym candidate generation model. The model was built using bayesian ranking.
- Developed tree-based gradient boosting model to predict out-of-stock items in supplier stores, leading to a 3–5% improvement in key inventory metrics across >100 stores in three countries.

FinAccel | The biggest BNPL startup in Indonesia. $1.6B USD Valuation. Jakarta, Indonesia
Data Scientist - Chatbot Feb 2020 – Feb 2021

- Developed email autoresponder and chatbot using LSTM models to support customer service operations, automatically resolving approximately 53% of daily incoming emails and 33% of daily chat inquiries.
- Built a product recommendation system to extract product embeddings and optimized the batch prediction speed by 8× using Approximate Nearest Neighbors (ANN) techniques on the product embeddings.

RESEARCH EXPERIENCE

Kata.ai | LLM and AI startup. Raised $3.6M USD in Series B. Jakarta, Indonesia
Research Scientist Sep 2018 – Sep 2019

- Published a paper comparing two methodology: Supervised Model Fine-Tuning and Unsupervised Pretrained LLM Fine-Tuning for Named Entity Recognition in Indonesian conversational texts.
- Finetuned the LLM on as few as 100 sentences and achieved an absolute improvement of 32 points in test F1 score compared to the baseline model.

i-machine-think | UvA and Meta AI Paris Research Group Amsterdam, The Netherlands
Graduate Research Student Jan 2018 – Aug 2018

- Published a paper on training neural networks to learn to follow instructions from small data, particularly in the subregular language and the language games of SHRDLURN.
- The study concludes that neural networks can automatically acquire the necessary inductive biases to learn to follow human instructions from minimal data.

Zylab | Software development company Amsterdam, The Netherlands
Research Scientist Intern Jul 2017 – Oct 2017

- Built their first CRF model for NER use case in English, Dutch, and French. Evaluated the CRF model on CoNLL 2003 dataset.

EDUCATION

Saarland University, Saarbruecken, Germany

- Ph.D in Computational Linguistics (Withdrawal, Unfinished)

University of Amsterdam, Amsterdam, The Netherlands

- M.Sc in Artificial Intelligence

University of Indonesia, Jakarta, Indonesia

- B.S. in Computer Science

PUBLICATIONS

[1] R. Leonandya, D. Hupkes, E. Bruni, and G. Kruszewski, "Training neural networks to learn to follow instructions from small data" in *Proceedings of the 13th International Conference on Computational Semantics (IWCS 2019)*. Gothenburg, Sweden, May 2019.

[2] R. Leonandya, F. Ikhwantri, "Pretrained language model transfer on neural named entity recognition in Indonesian conversational texts" in *The 33rd Pacific Asia Conference on Language, Information and Computation (PACLIC 33)*. Hakodate, Japan, Sep 2019.

[3] R. Leonandya, B.D. Trisedya, and N.H. Praptono, "A semi-supervised algorithm for Indonesian named entity recognition" in *Proceedings of the 3rd International Symposium on Computational and Business Intelligence*. Bali, Indonesia, Dec 2015.

TOOLS

Python, SQL, Docker, Jupyter Notebook, Flyte (workflow orchestration), Merlin (kubernetes for ML deployment), Apache Kafka, Cookiecutter, PyTorch, Tensorflow, Google Cloud Platform, Java, C#, LLMs.

LANGUAGES

Indonesian (native), English (fluent)