Reza Dwi Utomo Data Scientist | ML Engineer | AI Enthusiast

Jakarta Metropolitan Area, Indonesia utomorezadwi@gmail.com utomoreza.github.io
linkedin.com/in/utomorezadwi utomorezadwi.medium.com github.com/utomoreza

Profile

An engineer specializing in data-driven analysis. Currently working as a Data Scientist at Telkom Indonesia and appointed as Head of Legal Data Analytics at Legal Analytics (Powered by Telkom Indonesia). Self-motivated AI enthusiast with 5 years of varied experience, such as in independent research in intelligent control, reliability in the manufacture, and machine learning.

Highly interested in tech and AI applications, e.g. intelligent transport systems, natural language processing, computer vision, machine learning, big data, data science, and other intelligent applications. Skilled in analytical programming, e.g. Python, MATLAB, and R, and in AWS Cloud environment. Having data science and machine learning certificates from Udacity, deeplearning.ai, Microsoft Certified, and IBM Professional Certificate, to name a few.

Skills

Python *(NumPy, Pandas, Scikit-learn, Matplotlib, Plotly, Dash, etc.)*, **Deep Learning** *(Keras, TensorFlow, PyTorch)*, **Natural Language Processing** *(NLTK, SpaCy, Gensim, FastText, Transformers (Hugging Face))*, **R** *(dplyr, ggplot, flexdashboard, shiny)*, **AWS** *(SageMaker, S3, EC2, API Gateway)*, **SQL** *(MySQL, MariaDB, SQLite)*, **NoSQL** *(Elasticsearch)*, **Git** *(GitHub, GitLab)*, **Linux Bash**, **MLOps** *(MLflow, DVC)*, **MATLAB** *(Simulink)*, **Social Network Analysis** *(Gephi)*, **Azure** *(Azure Machine Learning)*, **Web Scrapping** *(BeautifulSoup, Selenium)*, **Deployment** *(Flask, FastAPI, Docker)*

Professional Experience

Head of Legal Data Analytics, Jan 2022 – present | Jakarta
Legal Analytics Powered by Telkom Indonesia
- Helping governmental and corporate organizations to implement Big Data and AI based law, legal, and social solutions.
- Translating requirements from users or management levels into Big Data and AI solutions.
- Delivering solutions for tens of billions of rupiahs worth of projects.
- Managing a team full of data scientists to develop and deliver Big Data and AI solutions.

Data Scientist, *PT Telkom Indonesia (Persero), Tbk.* Oct 2020 – present | Jakarta
- Involved in a data science team to work on the project to implement Natural Language Processing for Indonesian law.
- Developed named entity recognition (NER) model for Indonesian law using spaCy, Keras, Tensorflow, and IndoBERT pre-trained model.
- Used Elasticsearch to analyze text data.
- Developed WhatsApp Bot and Telegram Bot.
- Analyzed public transport data (ticketing) for gaining insights.
- Developed a methodology to categorize potential risks in social media and news texts.
- Developing Text Summarisation to summarise Registrasi docs of Mahkamah Konstitusi (Applications of Indonesian Constitutional Court).
- Developed BERT-based Text Similarity.
- Working on online news analytics.
- Working on social media analytics.
- Developing Affective Text Generation.
- Managing model deployments.

Data Scientist, *CODEX Powered by Telkom Indonesia* Apr 2020 – Sep 2020 | Jakarta
- Involved in a data science team to work in the project to implement Natural Language Processing for Text Similarity in law.
- Involved in a voluntary data science team to work on the project of the COVID-19 PeduliLindungi App . Scrapped COVID-19 data from various official websites of regional governments of Indonesia.
- Implemented Anomaly Detection in the project of big data analytics at PT Pupuk Indonesia.

Engineer, *PT Industri Kereta Api (Persero)* Feb 2019 – Dec 2019 | Madiun
- Performed statistics-based (e.g. descriptive, inferential, regression) reliability prediction analysis (developed in MATLAB) on passenger coach whose results are used as a base for managerial decision.
- Assigned to develop a dedicated-for-INKA concept of work for RAMS analysis (Reliability, Availability, Maintainability and Safety) of the rolling stock system.
- As the main contributor to the development of RAMS management process at the company.
- Performed the RAMS management Process using EN 50126.

Junior Expert, May 2018 – Jan 2019 | South Tangerang
Agency for the Assessment and Application of Technology (BPPT)
- Performed statistics-based (e.g. descriptive, inferential, regression) reliability prediction analysis.
- Developed FTA (Fault Tree Analysis) in R using RStudio.
- Involved in the LRT Greater Jakarta project.
- Worked in an engineering team collaborated with INKA Ltd.
- Performed RAMS analysis (e.g. FTA, RBD, FMEA, MTTR, MTBF, EN 50126) for the systems of the doors, HVAC, bogie, wiring and piping connections, and control panels of the rolling stocks.

Education

MRes in Railway Systems Integration, *University of Birmingham* Aug 2018 – Sep 2018 | Birmingham, UK
- Core modules: Mathematics as an Engineering Tool, Railway Operations and Control Systems, Railway Traction Systems Design, Railway Control Systems Engineering, Research Skills and Research Environment.
- The thesis topic is to develop ATO (Automatic Train Operation) Control Systems with Kalman filter (developed in MATLAB) with a case study of Docklands Light Railway in London. A part of the thesis was presented at the IEEE International Conference on Intelligent Rail Transportation (ICIRT) December 2018 in Singapore and published on IEEExplore .

BEng in Computer Engineering, *Diponegoro University* Sep 2010 – Aug 2015 | Semarang
- Core modules: artificial intelligence, fuzzy logic, neural networks, real-time operating system (RTOS), microprocessor design, embedded system, and distributed embedded system.
- The final project topic was to develop a fuzzy logic controller (developed in MATLAB/Simulink) able to control the transfer function of the train—published in IEEE Conference Proceedings .

Courses

MLE for Production (MLOps) Specialization, *deeplearning.ai* Jul 2022 – present
The Machine Learning Engineering for Production (MLOps) Specialization covers how to conceptualize, build, and maintain integrated systems that continuously operate in production. In striking contrast with standard machine learning modeling, production systems need to handle relentless evolving data.

Data Scientist Nanodegree, *Udacity* Aug 2021 – Dec 2021
The student will master the skills necessary to become a successful Data Scientist. He will work on projects designed by industry experts, and learn to run data pipelines, design experiments, build recommendation systems, and deploy solutions to the cloud.

Natural Language Processing Specialization, *deeplearning.ai* Mar 2021 – Jul 2021
The student will be ready to design various NLP applications. This Specialization is designed and taught by two experts: Younes Bensouda Mourri (an Instructor of AI at Stanford University) and Łukasz Kaiser (a Staff Research Scientist at Google Brain and the co-author of Tensorflow, the Tensor2Tensor and Trax libraries, and the Transformer paper).

TensorFlow Developer Professional Certificate, *deeplearning.ai* Dec 2020 – Jan 2021
The student will learn the necessary tools to build scalable AI-powered applications with TensorFlow. After finishing this program, he will be able to apply his new TensorFlow skills to a wide range of problems and projects. This program can help you prepare for the Google TensorFlow Certificate exam and bring you one step closer to achieving the Google TensorFlow Certificate.

Deep Learning Specialization, *deeplearning.ai* Sep 2020 – Dec 2020
The student will build and train neural network architectures such as Convolutional Neural Networks, Recurrent Neural Networks, LSTMs, Transformers, and learn how to make them better with strategies such as Dropout, BatchNorm, Xavier/He initialization, and more.

IBM Data Science Professional Certificate, *Coursera* Nov 2019 – May 2020
The program consists of 9 online courses that will provide the student with the latest job-ready tools and skills, including open-source tools and libraries, Python, databases, SQL, data visualization, data analysis, statistical analysis, predictive modeling, and machine learning algorithms. He will learn data science through hands-on practice in the IBM Cloud using real data science tools and real-world data sets.

Data Science Academy, *Algoritma Data Science Academy* Jan 2020 – Mar 2020 | Jakarta
The curriculum consists of two specializations: Data Visualization and Machine Learning . The former is a 4-week bundle of courses curated to accelerate a student's mastery in building data products, cleaning data, visualizing data in interactive dashboards, and extracting actionable insights. The latter is an 8-week bundle curated to accelerate a student's mastery in building data products, developing machine learning models, and understanding the core components that form most of industrial-grade AI today.

Certificates

- Microsoft Certified: Azure Fundamentals
- Microsoft Certified: Azure AI Fundamentals