Fathinah Asma Izzati

fathinah.izzati@mbzuai.ac.ae | Linkedin | Google Scholar | Github | +971 56 937 6401 | Abu Dhabi, UAE

Education

M.Sc. in Machine Learning, Mohamed bin Zayed University of AI, UAE   Aug 2023 – May 2025
CGPA: 3.95/4.0
Full-Merit Scholarship
Research Areas: Multimodal LLM, Controllable Generative Models (Music/Audio/Visual)
B.Sc. in Computer Science, University of Indonesia   Aug 2018 – Jul 2022
CGPA: 3.6/4.0
Achievements: The Most Outstanding Student, Stanford University Innovation Fellow, Start-Up Founder
Exchange Student, School of Computer Science, University of Birmingham, UK   Jan 2020 – Jun 2020
Erasmus+ Scholarship
CGPA: First-Class Honors

Experience

Instructor, Dept. Computer Science, Higher Colleges of Technology, Abu Dhabi, UAE   Jun 2025 – present
AI Research Scientist Intern, Core 42, Abu Dhabi, UAE   Aug 2024 – Oct 2024
- Explored novel architecture designs for multilingual speech-to-speech conversational systems
- Researched benchmarking to evaluate performance of different LLM architectures
- Developed training strategies and data preparation techniques to optimize diverse LLM architectures

Data Analyst Intern, First Abu Dhabi Bank, Dubai, UAE   May 2024 – Aug 2024
- Analyzed HR data (hiring, recruitment, performance, events) across the entire bank, identifying key insights and visualizing them in Power BI
- Automated data cleaning and daily reporting workflows from multiple sources (Smart Recruiter) into Power BI
- Automated the hiring events pipeline by creating a queue system and website for candidates and recruiters

NLP Engineer, Willowmore Pte Ltd, Singapore   Jun 2023 – Oct 2023
- Developed AI chatbot using open-source LLMs (GPTs, LLaMA, BERT, Vicuna, Guanaco)
- Implemented zero-shot learning with query expansion from vector databases to enhance QA capabilities.
- Deployed over 10 LLMs on edge devices

Data Scientist (Risk), Advance.AI, Indonesia   May 2022 – May 2023
- Developed high-AUC credit risk models using logistic regression and gradient boosting.
- Engineered over 150 features using PostgreSQL; optimized deployment queries by 50%.
- Conducted comprehensive market analysis on Indonesia's financial penetration (loan usage, default rates, etc.), delivering actionable insights to stakeholders
- Assisted a collaborative project with Indonesia's largest telco company, generating over 500 features from telecom data
- Automated score updates using cron jobs and bash scripts; performed large-scale data processing

Data Scientist Intern, PT XL Axiata Tbk, Indonesia   Aug 2021 – Feb 2022
- Developed an end-to-end ML pipeline to predict telecom tower revenue
- Improved model accuracy by 24% and reduced RMSE by 9.9%

Data Engineer Intern, OVO (PT Visionet Internasional), Indonesia   Oct 2020 – Apr 2021
- Developed ML-based anomaly detection system for ETL monitoring
- Engineered data pipelines integrating PostgreSQL, MongoDB, and APIs into a data warehouse
- Managed data warehouse operations and facilitated ad-hoc data ingestion for multiple divisions

Awards

1st Winner of National AI Hackathon, COMPFEST   2022
The Most Outstanding Student, Department of Computer Science, University of Indonesia   2021
Google Bangkit ML Graduate and Best Project, Google, Indonesia   2021
2nd Runner-Up of Startup Games, University of Birmingham   2020
34K readers in Medium and featured in the top reading list,
I wrote articles about software design, development, and deployment (fathinah.medium.com)   2020

Featured In

- University of Indonesia News
- Antara News

- University Innovation Fellows Profile

Research

EXPOTION: Facial Expression and Motion Control for Multimodal Music Generation Mar 2025
- Accepted in IEEE ISMIR 2025

Leveraging MU-LLaMA for Scene Imagination and Enhanced Video Background Music Generation May 2024
- Constructed a large-scale video-audio caption dataset using MU-LLaMA (MuBERT + LLaMA-2).
- Enhanced MusicGen by integrating a music captioning module; fine-tuned MU-LLaMA for cross-modal tasks.

Spoken Dialects to Modern Standard Arabic Translation May 2024
- Developed speech-to-speech conversion from dialects to MSA by fine-tuning models like HuBERT, Whisper, and Wav2Vec.
- Built cascaded architecture consisting of ASR, translation (T5), and TTS modules (MMS-TTS).

Integrated Statistical and Predictive Modeling for Alzheimer's Disease May 2024
- Developed predictive models using neuroimaging and multi-omic data; achieved up to 98% accuracy.
- Employed ensemble models and Graph Neural Networks; conducted high-dimensional data analysis.

A Comparative Analysis of LLMs and Embeddings for MCQ Answering Dec 2023
- Examined LLMs and embeddings for MCQ answering; compared models like Platypus2, GPT-3.5, BERT, and RoBERTa.
- Enhanced performance with Wikipedia RAG; evaluated using MAP@3.

Note: Some of the projects above are submitted to conferences and are currently under review.

Curated Projects

Deep Music Generation, Presentation Mar 2024
- Leveraging prior sampling, chord and texture recombination, and interpolation techniques to create dynamic, expressive compositions. Utilized generative models like EC2-VAE and Accomontage.

Algorithmic Music Composition, Presentation Jan 2024
- Engineered pop music compositions using advanced synthesis techniques, including AM/FM synthesis, ADSR envelopes, Bytebeat, and granular synthesis.

Predicting US Stocks Closing Movements Dec 2023
- Conducted AI research to predict Nasdaq stock closing prices utilizing models like LightGBM, ARIMA, GARCH, and LSTM.
- Implemented extensive feature engineering and time series analysis to enhance financial analytics and stock prediction methodologies.

E-KIPI: Smart Monitoring System of COVID-19 Vaccine's Adverse Effect Jul 2021
- Created a mobile app to report COVID-19 vaccine adverse effects and give users relevant advice from text input
- Built models to predict side effect severity from text and face recognition model
- Deployed models on-cloud, created API, and integrated with mobile app
- Tech-stack: Android Kotlin, Tensorflow Keras API, Google Cloud Platform, REST API

Road Segmentation using Deep Learning UNet and Conventional Model Dec 2020
- Extracted the road network from Massachusetts roads dataset images by using a deep-learning model called U-Net and a non-deep learning model (image segmentation and morphological operations)
- I applied histogram thresholding, morphological operation (opening), and region-based segmentation and achieved 0.105 IoU score

Teaching Experience

Teaching Assistant, Discrete Mathematics (CSGE601010) Fall 2021
Teaching Assistant, Database (CSGE602070) Spring 2021
Teaching Assistant, Programming (CSGE601021) Fall 2020
Speaker, Open House, Computer Science University of Indonesia 2024
Speaker, Business IT Case Competition, COMPFEST 2024
Speaker, Data Science Academy, COMPFEST 2023
Speaker, Data Science Workshop, Educare Fasilkom UI 2023
Speaker, Data Analysis Workshop, Generation Girl 2023
Mentor, Data Analytics Workshop, RISTEK Fasilkom UI 2022

Organization Experiences

University Innovation Fellow, Stanford University's Hasso Plattner Institute of Design (d.school) Aug 2021 – Jul 2022

- Selected and funded by Indonesia's Ministry of Research and Technology to undergo a 12-month intensive training in design thinking and lean startup methodologies.
- Developed the first 'Peer-Based CV Review' web platform for Universitas Indonesia students

Senior Associate of Data Science SIG, RISTEK Fasilkom UI   Feb 2021 – Jan 2022
- Taught members on Data Analytics, Machine Learning, and Deep Learning concepts through weekly internal classes.
- Contributed in backend development using FastAPI for the recommendation system model, PANSOS application: Repository, News Article, Official Page

Founder, Reviewin.me   Jan 2022 – Dec 2022
- Launched a non-profit startup offering peer-review services for students' CVs and resumes.
- Managed a team of 10, overseeing the tech and design departments.
- Led activities including prototyping, user research, application development (frontend and backend), usability testing, load testing, and user acceptance testing.
- Developed the frontend using NEXT.js and integrated it with a GraphQL backend.

Skills

Programming: Python, Java, SQL, Bash, Unix/Linux
Frameworks: PyTorch, TensorFlow, HuggingFace, Fine-Tuning, FastAPI, OpenAPI, Distributed Training, RAG
Tools: GCP, Docker, Airflow, Hadoop, Hive, MongoDB, PostgreSQL, Git, PowerBI, Tableau
Certifications:
- TensorFlow Developer Certificate
- DeepLearning.AI TensorFlow Developer
- Mathematics for Machine Learning
- Google IT Automation with Python