RAFI MADANI

rafimadani78@gmail.com | 088975952107 | https://www.linkedin.com/in/rafi-madani-320822218/

EXECUTIVE SUMMARY

Computer Science student (Universitas Indonesia) passionate about NLP, LLMs, and Generative AI. Developed AI solutions using DeepSeek, LLaMA, Mistral, and Qwen models with advanced prompting strategies including zero-shot and few-shot learning. Delivered production-grade systems across internships and multiple capstone projects. Currently researching LLM-based content classification for thesis work. Strong foundation in Python, Django, and ML frameworks, with interest in RAG architectures, VectorDBs, and agent-based systems. Adept at deploying AI and ML models into production environments via Flask and Streamlit. Passionate about scalable, ethical, and real-world AI applications.

EDUCATION

Computer Science Bachelor, Universitas Indonesia April 2021 - Now

- Undergraduate Thesis : Perbandingan Performa Large Language Model (LLM) dalam Klasifikasi Konten Pornografi pada *Tweet* Berbahasa Indonesia
    - Engineered a content moderation system using state-of-the-art LLMs (LLaMA 3.1 8B, DeepSeek, Qwen, Mistral)
    - Applied prompt-based few-shot and zero-shot classification (no fine-tuning)
    - Built tools for misclassification analysis using log-probabilities and token-level confidence scores
    - Identified and mitigated model bias using curated few-shot examples
- Related coursework: Calculus, Linear Algebra, Statistic and Probability, Databases, Introduction to Artificial Intelligence & Data Science, Image Processing, Numerical Analysis, Django, Backend, OOP, Python

BANGKIT ACADEMY LED BY GOOGLE, TOKOPEDIA, GOJEK, & TRAVELOKA Feb 2024 - June 2024
Machine Learning Cohort

- (Capstone) Collaborated with a team of seven people and successfully selected as the sole team out of 103 teams to undertake a mobile app project commissioned by Nusantara Beta Studio, a renowned company known for developing widely used mobile applications.
- Best Capstone Presenters (Individual - 1 per team)

SKILLS

Languages : Python, HTML/CSS, Kotlin
Backend: PostgreSQL, Django, JavaScript Libraries, Object-Relational Mapping (ORM), CI/CD, TDD
Machine Learning: Tensorflow, Keras, Convolutional Neural Networks (CNN), Natural Language Processing (NLP), Deep Learning, Computer Vision, Model Deployment, Regression, Recommender Systems, NumPy, Pandas, Streamlit
Large Language Models (LLMs): DeepSeek, LLaMA, Mistral, Qwen
Prompt Engineering: In-Context Learning (ICL), Chain of Thought (CoT), Zero-shot & Few-shot Prompting

EXPERIENCE

PT. PHAROS September 2024 - December 2024
Software Engineering Intern

- Developed and optimized existing web-based system using company's framework, Toraja framework, including a sales dashboard and automated incentive calculations. Improved query performance, reducing loading times by 97% and query execution by 99%
- Designed and implemented WhatsApp Blast functionality for clinic owners and doctors, utilized by approximately 100 field force personnel to engage with around 900 clinic recipients. The development encompassed both the frontend (Toraja Page) and backend (Django ORM) systems
- Built a comprehensive WhatsApp notification system with multi-PDF delivery, targeted messaging, and retry mechanisms. Optimized performance via caching and prevented throttling.
- Created a PDF generator for doctor certificates (SKP) with dynamic content, auto-fill, and batch processing. Improved generation speed using caching and worker processes.

PROJECTS

Movie Recommendation App, Nusantara Beta Studio Company April 2024 – June 2024

- Collaborated with a team of three to develop a machine learning model for a movie recommendation app, deployed on mobile platforms.
- Conducted data scraping using the TMDB API to gather approximately 2500 clean movie data.
- Used TF-IDF vectorization to transform movie overviews into numerical representations, and computed cosine similarity between these vectors to measure the similarity between movies.
- Developed PyTorch-based neural collaborative filtering model achieving a 82% confidence result with hit and ratio.
- Successfully deployed a web-based movie recommendation service using Flask.

Logistic Management Website, RS UMMI BOGOR February 2024 – May 2024

- I led a project to develop a logistic management system website for RS Ummi, a hospital in Bogor. Utilizing Python for the back-end, Django REST as the framework, Tailwind for the front-end, Scrum as a methodology, and Supabase as a main database.
- Played a dual role as both the Scrum Master and the Lead Back-End Developer. As a result, my project's code coverage always exceeds the definition of done, which is 90% with the practice of TDD.
- Delivered the project at the exhibitions and attracted approximately 25 visitors

Sarcasm Detector Website, Personal Project May 2024 – June 2024

- Developed and deployed a Natural Language Processing (NLP) model to detect sarcasm in news headlines, achieving over 80% validation accuracy.
- Implemented data preprocessing techniques, including tokenization and padding, to handle text data efficiently.
- Enabled automatic title generation and sarcasm prediction by inputting news article links, enhancing user convenience and functionality.
- Deployed the application using Flask, allowing users to interact with the model via a simple web interface.

Bike Sharing Analysis, Dicoding Feb 2024 – March 2024

- Analyzed how weather and seasons affect bike rental numbers to help a bike rental business owner make informed decisions.
- Created visual reports with barchart to present findings on rental trends based on weather and seasons.
- Deployed the findings using Streamlit for easy access and website support.

Flight Delay Data Analysis, Universitas Indonesia Oct 2023 – Dec 2023

- Identified patterns and trends in flight delays across different days of the week, aiding in the understanding of delay causes and visualized the results using a stacked bar chart, and potentially improving scheduling and operational strategies.
- Developed a machine learning model to predict the ground time of airplanes for each flight using a Multilayer Perceptron (MLP) regressor. Achieved a Mean Squared Error (MSE) of 64.867 and an R-squared (R^2) of 0.848, demonstrating the model's effectiveness in predicting ground time.

- Visualized the top 10 routes using a horizontal bar plot with distinct colors for each airline.

ORGANIZATION

Forum Ukhuwah dan Kajian Islam Fasilkom UI Mar 2023 - Jan 2024
Head Division Of Social And Politic
- Managed and edited infographics on social issues, achieving approximately 300 likes across 2 posts.
- Organized a sacrificial animal donation event, successfully arranging for the donation of one goat for the campus community.
- Coordinated weekly Friday charity drives, raising a total of IDR 1.5 million.

Your Best Shelf Indonesia Jul 2022 - Feb 2023
Head Division Of Literacy Development
- Hosted events like webinars and book discussions to improve literacy, with one meeting attracting 20 participants.
- Strategically managed and utilized the division's human resources to align with our vision and fulfill our mission, ensuring efficient operation and impactful literacy initiatives.
- Developed social media content to raise literacy awareness, achieving 541 likes across three posts.

CERTIFICATION

- **Test Of English As Foreign Language (ITP) : 587**
- **DeepLearning.AI TensorFlow Developer by DeepLearning.AI on Coursera** | Certificate
- **Machine Learning by Stanford University & DeepLearning.AI on Coursera** | Certificate
- **TensorFlow: Advanced Techniques by DeepLearning.AI on Coursera** | Certificate
- **TensorFlow: Data and Deployment by DeepLearning.AI on Coursera** | Certificate
- **Mathematics for Machine Learning and Data Science by DeepLearning.AI** | Certificate