Adam Nguyen

Date of Birth: 7th of May, 2001
Nationality: Vietnamese
Address: Tan Phuoc Apartment, Ly Thuong Kiet Street, District 11, Ho Chi Minh City, Vietnam

Phone: (+84) 923369122
Email: AdamaraNguyen@gmail.com
Github: @duonggiakhanhb
LinkedIn: @adamaranguyen

As an ICPC contestant, I excel in solving complex algorithms. I have experience in writing SQL and NoSQL queries, utilizing tools in Data Engineering, and proficiently working with AWS services.

WORK EXPERIENCE

HAHALOLO, **Ho Chi Minh, Vietnam**
AI Engineer *Jun 2022–Jun 2023*

- Processed and refined data to optimize AI model performance.
- Built ETL pipelines for efficient data extraction, transformation, and loading.
- Skillfully extracted data from online sources for analysis.
- Conducted research, enhanced methodologies, and deployed impactful ML models for data-driven decisions.

EDUCATION

HO CHI MINH UNIVERSITY OF EDUCATION **Ho Chi Minh, Vietnam**
Bachelor of Computer Science *September 2019–Aug 2023*
Honors: cum laude (GPA: 3.55/4.0)

AWARDS

Student Scientific Research Awards 3rd Prize: Applied Machine Learning and Knowledge Inference to design MathSolver Scanner–a Chatbot System for answering basic handwritten equation.

SKILLS

- **Programming and Scripting**: Proficient in **Python**, **JavaScript**, **Scala**, **C++**, and **UNIX**. Experienced in using **Tensor**, **Torch**, **Pandas**, **NumPy**, and other data manipulation libraries.
- **Data Processing and Analytics**: Skilled in using **Hadoop**, **Spark**, **Kafka**, **Apache NiFi**, **Airflow**, **Iceberg**, **Apache Superset**, **Tableau**, and **Power BI** for ETL, BI, big data, and machine learning projects.
- **Database and Data Storage**: Familiar with **Cassandra**, **PostgreSQL**, **MySQL**, **MongoDB**, **SQLite3**, **Minio**, **Nessie Catalog**, and **AWS S3** for data warehousing, data modeling, SQL, and NoSQL operations.
- **Cloud Computing and AWS**: Proficient in **AWS S3**, **Athena**, **Glue**, **Redshift**, **QuickSight**, **Kinesis**, **CloudWatch**, **Lambda**, and **EMR** for cloud computing, cloud architecture, and cloud security.

- **Web Development and Backend**: Experienced in **HTML**, **CSS**, **Web Component**, **React**, **Django**, **Laravel**, **FastAPI**, and **Django REST Framework** for web development, RESTful APIs, and microservices.
- **Problem Solving and Soft Skills**: Analytical mindset, quick learner, passion for practical data engineering architecture, basic English proficiency. Demonstrated ability to solve complex data problems and deliver high-quality results in fast-paced environments.

PROJECTS

HAHALOLO

USER INTERACTION ETL PIPELINE DEVELOPMENT

Contributed to the construction of an ETL pipeline to extract, transform, and load user interaction data into the central data warehouse. Collaborated within a team to design and optimize the pipeline, ensuring seamless integration and accurate data representation.

IMPROVED TEXT VIOLATION DETECTION MODEL

Enhanced the text violation detection model, elevating accuracy from 67.2% to 87.8%. Implemented refined methodologies and fine-tuned algorithms, leveraging advanced techniques to better discern and categorize inappropriate language.

FRIEND RECOMMENDATION ALGORITHM ENHANCEMENT

Participated in enhancing the friend recommendation algorithm, utilizing behavioral patterns and user preferences to improve suggestions. Fine-tuned the algorithm to offer more precise and relevant friend recommendations based on user interactions.

HOTEL AND TOUR RECOMMENDATION SYSTEM

Developed a recommendation system for hotels and tours, analyzing customer behaviors and preferences to build an algorithm for personalized recommendations. Optimized the algorithm to deliver effective suggestions, enhancing customer experience and value proposition.

MY OWN PROJECT

MOTOGP DATA ENGINEERING PROJECT [Link to Github]

Developing a comprehensive data engineering project centered around MotoGP racing. Collecting, processing, and visualizing MotoGP data to gain insights into riders' performance, team statistics, race outcomes, and more. Leveraging Apache Airflow, Apache Kafka, Apache Spark, and Cassandra for data processing and storage.

SIMPLE MINI DATA LAKEHOUSE ON YOUR LOCAL [Link to Github]

Developed a local data lakehouse for synthetic banking data using Python, Apache Spark, and related tools. Generated synthetic data with Python Faker, loaded it into MinIO, and utilized Dremio for cataloging, querying, and analytics. Integrated Nessie Catalog and Apache Iceberg for versioning and metadata management. Analyzed and visualized data through Superset.