Phuong-Dung NGUYEN

Available from September 2024
Paris, France
dzungnguyen.contact@gmail.com ♦ (+33) 6 66 14 66 64
dzungpnguyen.github.io ♦ [linkedin/dzungnguyen00](https://linkedin.com/in/dzungnguyen00)

Data Scientist / Data Engineer

Recent graduate with strong background in Machine Learning/AI, I'm seeking a CDI opportunity in the data universe. Having 2 years of experience as Data Scientist/Data Engineer, I'm proficient in data manipulation, developing ML-based models and data pipelines. Motivated by challenging missions that go beyond Python and SQL, where I will contribute meaningfully while learning continuously.

"I don't need an alarm clock, my code wakes me up with errors."

EDUCATION

Master of Engineering in Big Data and Machine Learning **September 2022 – September 2024**
EFREI Paris, Villejuif, France
♦ Data Science, ML/AI, Deep Learning, MLOps, Data Engineering, DevOps, Database Management

Bachelor Data Science and Statistics **September 2021 – September 2022**
Lumiere Lyon II University, Lyon, France
♦ Artificial Intelligence, Data Science with Python, Statistics, Data Structures and Algorithms

CORE COMPETENCIES

Model development ♦ Algorithm Study ♦ DevOps ♦ MLOps ♦ Data Analysis ♦ Data Visualization ♦ Database Manipulation ♦ Programming ♦ Version Control (Git/GitHub) ♦ Problem-solving ♦ Project Management ♦ Communication ♦ Teamwork

Data Science: R ♦ NLP ♦ Computer Vision ♦ Python (Pandas, Numpy, Scikit-learn, TensorFlow, PyTorch, Streamlit, Plotly, Seaborn…)
Data Engineering: Data Pipeline ♦ Docker ♦ GitHub Actions ♦ CI/CD ♦ Java ♦ Scala ♦ Spark ♦ Apache Kafka ♦ Azure
Database Management: RDBMS (MySQL, PostgreSQL, Oracle) ♦ NoSQL DB (MongoDB, Redis, Cassandra, Neo4J)
Web Development: Java ♦ SpringBoot ♦ Flask ♦ Django

EXPERIENCE AND ACHIEVEMENTS

Data Scientist - Data Engineer Apprentice **September 2022 – September 2024**
Societe Generale, Paris, France
♦ Conduct data exploration and data analysis to gain insights into the data model.
♦ Develop models to assess client companies' environmental impact, focusing on CO_2 emissions. Achieved an accuracy rate of 97%.
♦ Build dada pipelines to retrieve and consolidate multi-format data from APIs, reducing 50% of processing time for 15,000 requests.

Software Engineer Apprentice **September 2022 – September 2024**
Societe Generale, Paris, France
♦ Collaborate in a team of 4, refactor a web application by introducing Restful API service and designing a statistic dashboard.
♦ Orchestrate UAT procedures and document the outcomes, identifying production environment errors before they impacted end-users.

PROJECTS

Model Comparison Image Classification **February 2024**
♦ Trained and compared 4 Deep-Learning-based models on a binary dataset for classification. The models include 1 without CNN layers, 1 with CNN layers and 2 transfer learning approaches with 2 Keras-based models.

Model Deployment via API on Cloud **February 2024**
♦ Implemented end-to-end DevOps practices to develop and deploy a KNN-based model, utilizing Docker for containerization and Azure for cloud deployment service while ensuring autoscaling through load testing.

Legal Chatbot **December 2023 – January 2024**
♦ Trained a ML-based chatbot capable to explain legal terms and resolve legal issues. This involved fine-tuning a T5 model using 4000 Q&A pairs scraped manually from forums of legal professionals.

Real-time Monitoring System **February 2023 – May 2023**
♦ Developed a Big Data solution in Scala with 3 side: Simulating drones that generate data every 3 seconds ; Monitoring data in real-time by Kafka for immediate alerting in critical situation ; Performing data analysis with Spark Analytics.

Data Analysis and Visualization: MovieLens and Estate Value in France **November 2022**
♦ Deployed 2 Streamlit applications for data analysis and data visualization (by Plotly and Seaborn) on datasets of over 30M records.

LANGUAGES

Vietnamese: Native
French: Fluent C1
English: Fluent C1 (955/990 TOEIC)