

TUAN NGUYEN

Data Scientist | Data Analyst

Address: 1300 30th Street, Boulder, CO | Phone: (303)-472-7172

Email: tuangn@colorado.edu | LinkedIn: www.linkedin.com/in/tuannguyen129

Personal Page: <https://tuangnuyen.me>

Data Scientist with 3+ years of experience in analytics, machine learning, and AI, specializing in chatbot solutions. Backed by 6 years in geophysics at a petroleum institute, focusing on seismic data analysis. Proficient in python, SQL, and graph databases, with a passion for applying data science to drive innovation and efficiency in the energy and other industries.

PROFESSIONAL EXPERIENCE

08/2022 – 07/2024 | Data Scientist | Vietnam Petroleum Institute (VPI) | Hanoi, Vietnam

- Developed a chatbot-operated map-based website that executed user commands for navigation, leveraging a graph database to link map functions with domain knowledge and provide real-time visual results besides chat responses.
- Developed an internal document-sharing web-based application integrated with a retrieval-augmented generation (RAG) chatbot system, significantly enhancing document sharing capabilities by enabling chatbot interaction with local domain data and optimizing for Vietnamese language.
- Developed an automated oil field daily reporting system, saving one employee's daily workload by streamlining data transfer to SQL database, summarization, and reporting (via LLM), while enabling multi-level management supervision.
- Optimized chatbot reasoning with domain knowledge by generating precise question-answer chains through prompt engineering, and customized integrated chatbots for Microsoft Teams using Automate, Copilot Studio, and Prompt Flow.
- Developed web application name VPI-Mlogs for applying data tools and machine learning algorithms to petrophysics borehole data including data loading, anomalies detection, data cleaning, data visualization, missing log prediction.
- Developed the "vpitools" Python PyPi package for petrophysics data processing, offering quick statistics, 2D/3D visualizations, and customized plotting using matplotlib, seaborn, and interactive charts with plotly and vega-altair.
- Processed well-log data: detected and removed outliers, duplicated values, normalizing data, feature engineering, optimized size of training and testing dataset, logged and visualized model performance.

07/2021 – 07/2022 | Data Analyst | Vietnam Petroleum Institute (VPI) | Hanoi, Vietnam

- Utilized Microsoft Power Virtual Agents and Power Pages to create domain question-answering system via web application and Microsoft Team agents.
- Constructed borehole relational database, queried, and visualized logs data in Power BI Dashboard.

09/2015 – 06/2021 | Geophysicist | Vietnam Petroleum Institute (VPI) | Hanoi, Vietnam

- Automatically and semi-automatically interpreted seismic data and built the environmental models to predict the location of potential unstructured oil and gas traps, supporting drilling decisions.
- Built a velocity model, generated seismic attributes and applied specialized techniques such as Amplitude versus Offset and Inversion, to identify potential field locations and calculate oil reserves.
- Interpreted seismic data and mapped objects to generate detailed subsurface maps, including structural and attribute maps, with a focus on reservoir characterization.

SKILLS

- Database: MySQL, Neo4j, MongoDB, Weaviate, Qdrant.
- Languages: Python (Advanced), HTML-CSS (Intermediate), JavaScript (Intermediate).
- Libraries: Scikit Learn, Spacy, Tensorflow, PyTorch, Numpy, Pandas.
- Visualization: Matplotlib, Seaborn, Plotly, Altair.
- Python web development: Streamlit, Gradio, Github, HuggingFace.

EDUCATION

- 08/2024 – 06/2026: Master Science of Data Science at University of Colorado Boulder.

- 09/2010 – 08/2015: Bachelor Science of Geophysical Engineering at Hanoi University of Mining and Geology, Vietnam.

CERTIFICATES & SHORT COURSES

- (2022) Data Scientist Associate (DataCamp Certification)
- (2020) VPI – Seismic Signal Processing, VPI – Advanced Seismic Data Interpretation, CGG – Geostatistical Inversion.

PUBLICATIONS

- (2022) VPI-Mlogs: A web-based machine learning solution for applications in petrophysics.
- (2021) Results of seismic surveys offshore the Northern part of Bac Bo Gulf.

RESEARCH & PROJECTS

09/2015 – 08/2024 | Vietnam Petroleum Institute (VPI) | Hanoi, Vietnam

- (2024) *Data Scientist* | Chatbot-driven map navigation with integrated graph database.
- (2024) *Data Scientist* | E&P knowledge sharing with RAG chatbot and knowledge graph integration.
- (2023) *Data Scientist* | Petroleum Journal chatbot system enhanced by LLM models.
- (2024) *Data Scientist* | VPI-Mlogs: Web-based application for petrophysical data analysis.
- (2023) *Data Scientist* | Daily Bach Ho oil field operations dashboard and reporting system using LLM.
- (2023) *Data Scientist* | Hydrocarbon composition analysis and oil product allocation.
- (2022) *Data collector, Data Analyst* | Well-log database for the Cuu Long basin area.
- (2022) *Data Analyst* | Fracture zones forecasting in the reservoir rock based on petrophysical logs.
- (2022) *Seismic modeler* | Provision service of Pre-stack inversion for DH01&02 Area - Block 05-1a.
- (2022) *Seismic modeler, Velocity Modeler* | Seismic Stratigraphy Analysis in Miocene and Oligocene sediments to determining unstructured trap at Block 09-1, Cuu Long Basin.
- (2021) *Seismic Data Analyst, Seismic Evaluation* | Merging of 3,000km² 3D PSDM Seismic Data from multi blocks, Cuu Long Basin, Vietnam.
- (2021) *Seismic Interpreter, Attributes Modeler* | Analysis, assessment of properties and distribution of Miocene reservoir in the Northern Bach Ho Field.
- (2020) *Data collector, Seismic Data Analyst* | Seismological interpretation of the central Song Hong basin focusing on the Upper Miocene-Pliocene sediments for oil and gas exploration and assessment.
- (2019) *Data collector, Seismic Data Analyst* | Database structuring for oil and gas open blocks to attract foreign investment in field exploration in Vietnam.
- (2018) *Seismic Attributes Modeler* | 3D PrSDM Seismic Interpretation in Block 09-3/12 and adjacent location of Block 09-1 (796km²), Cu Long basin, Offshore Vietnam.
- (2017) *Seismic Interpreter, Mapper* | 3D Petroleum System Basin Modeling Study, Block 06.1, Nam Con Son Basin, Offshore Vietnam.
- (2016) *Seismic Interpreter, Mapper* | Structural Interpretation of PPPSTM 3D/4C Seismic Data (847 km²), Block 09/1, Cu Long basin, Offshore Vietnam.
- (2015) *Seismic Interpreter* | Research on the formation and accumulation of oil and gas in the Miocene sediments of the central Song Hong basin, Offshore Vietnam.