

# LE QUOC AN

## DATA SCIENTIST INTERN

**LinkedIn:** [linkedin.com/in/anquoc071](https://www.linkedin.com/in/anquoc071)  
**GitHub:** [github.com/quyda2004](https://github.com/quyda2004)

**Email:** [anquoc071@gmail.com](mailto:anquoc071@gmail.com)  
**Mobile:** (+84) 398055256

### EDUCATION

---

**University of Science VNU-HCMC**  
Major of Data Science: GPA: 3.55/4.0

Ho Chi Minh City, Vietnam  
2022 - 2026 (Expected)

### WORK EXPERIENCE

---

#### • AI ENGINEER INTERN

*Wisdom Engineering and Business Solutions*

**01/2025 – 06/2025**

- Explored and implemented Reinforcement Learning algorithms (Q-Learning, Deep Q-Network) for adaptive traffic distribution in SDN environments.
- Designed and deployed network topologies in Mininet and configured Ryu Controller for SDN-based simulations.
- Built a custom traffic generator and data collection framework to simulate real network loads.
- Applied and benchmarked traditional load balancing algorithms (Round Robin, Least Connections) against learning-based methods.
- Developed and trained a DQN Agent to optimize decision-making based on live network data.
- Integrated graph-based network representations to enhance RL agent performance.

### PROJECTS

---

#### **Fake News Detection (Apr 2025 – May 2025)**

[\[GitHub\]](#)

**Team Size:** 3 — **Role:** AI Engineer/ Data Scientist (Team Leader)

- Built fake news dataset (FAKE / TRUE).
- Trained traditional models: TF-IDF / CountVectorizer + Logistic Regression, SVM, Random Forest, Decision Tree, Naive Bayes.
- Fine-tuned deep learning models: BERT, XLNet.
- Achieved 99.33% accuracy with BERT/XLNet.

#### **House Price Prediction (July 2024 – Aug 2024)**

[\[GitHub\]](#)

**Team Size:** 4 members — **Role:** Data Scientist/ Data Analyst (Team Member)

- Scraped real estate data and extracted key features using LLMs via Anyscale API.
- Preprocessed the dataset and conducted exploratory data analysis.
- Developed predictive models and compiled a report detailing insights and feature analysis.

#### **EEG-Based Mental Attention State Classification (Nov 2024 – Dec 2024)**

[\[GitHub\]](#)

**Team Size:** 4 members — **Role:** Data Scientist (Team Leader)

- Processed raw EEG data and built classification models (SVM, Random Forest, CNN) to identify mental states.
- Performed hyperparameter tuning to optimize classification performance.

#### **RAG-based AI Chatbot – Semantic Search & Gemini QA (May 2025 – Jun 2025)** [\[GitHub\]](#)

**Team Size:** 1 — **Role:** AI Engineer/Data Scientist

- Developed a Retrieval-Augmented Generation (RAG) chatbot combining SentenceTransformer and Gemini 1.5 Flash.
- Crawled and preprocessed website content; generated embeddings using `all-mpnet-base-v2`.
- Retrieved relevant contexts via semantic search (cosine similarity).
- Generated context-aware answers using Google Gemini API.

## SKILLS SUMMARY

---

- **Programming:** Python (NumPy, Pandas, Scikit-learn, PyTorch, Flask/REST API), C/C++, SQL, NoSQL, R
- **Data & ML:** Data preprocessing, visualization (Matplotlib, Seaborn, Plotly), Data Mining, ML/DL (CNN, RNN, LSTM, GRU, XGBoost, Random Forest, SVM,...)
- **NLP:** Chatbot development, sentiment analysis, text classification, NLTK
- **LLM:** Working with Transformer-based models (BERT, GPT, XLNet), fine-tuning, Retrieval-Augmented Generation (RAG)
- **Model Tuning:** Grid Search, Random Search, Bayesian Optimization
- **Tools:** Git, GitHub, MySQL, PostgreSQL, MongoDB, Docker, Postman
- **Soft Skills:** Strong teamwork, independent working, self-motivated
- **English:** B2 (Reading)