

# Hiep Tan To

Ho Chi Minh City – Viet Nam

☎ +84 919703340 • ✉ tanhiep2012003@gmail.com

Bachelor's and Master's in Computer Science student at Ho Chi Minh University of Science, specializing in Software Engineering, Backend Development, and System Integration. Strong foundation in OOP, algorithms, and database systems. Experienced in building scalable pipelines, system-level applications, and contributing to open-source software.

## Education

<b>Master of Computer Science</b>	In Progress
University of Science - VNUHCM	2025
<b>Bachelor of Technology with Honors</b>	9.41/10.0
Computer Science and Engineering, University of Science - VNUHCM	2021-2025
<b>Thesis</b>	10.0/10.0
Generative AI-Based Virtual KOL System	2025
<b>Language Skills</b>	IELTS 6.0, VSTEP B2

## Internships and Experience

### Junior AI Engineer

*First AI*

*Jul. 2025 - Present*

Developed recommendation systems for AI-driven search and social platforms, improving content relevance and user engagement.

Trained and fine-tuned machine learning and LLM-based models using user behavior and contextual data.

Implemented retrieval-augmented generation (RAG) to enhance recommendation accuracy with external knowledge.

Integrated external LLM APIs to support semantic search and personalized recommendations.

Optimized model inference for real-time recommendation services in production.

Evaluated recommendation performance using metrics such as Precision@K, Recall@K, and NDCG.

### Associate Machine Learning Engineer

*DeepCDR Team, OPSWAT*

*Jul. 2025 - Jul. 2025*

Developed AI for file type detection, optimizing accuracy from 85% to 92%.

Designed and trained classification models on large-scale file datasets, improving detection accuracy.

Evaluated model performance using precision, recall, and F1-score metrics across various file types.

Collaborated with cybersecurity experts to align ML models with threat detection requirements.

Optimized model inference for integration into production pipeline with minimal latency.

Conducted research and benchmarking on traditional and deep learning approaches for file analysis.

### Contractor Software Engineering - Treasure Data

*Integration Team, Treasure Data*

*Apr 2024–Jun 2025*

Developing and maintaining connectors for the Integration team using Embulk with Treasure Data solutions.

Collaborating with cross-functional teams to design and optimize data pipelines for robust migration and consistency, implementing and refining ETL processes, troubleshooting issues, and enhancing system performance.

Streamlining data integration workflows to support business intelligence initiatives.

Assisting in developing a new AI agent for summarizing documents.

Improved the application to facilitate CLI manual testing for QA.

### Undergraduate Thesis: Generative AI-Based Virtual KOL System

*Guide: Dr. Le Trung Nghia, Assoc.Prof. Tran Minh Triet, University Of Science*

*Sept. 2025*

Developed a generative AI system merging identity with varied styles, preserving semantics.

Built a modular, extensible architecture to integrate new algorithms and custom pipelines easily.

## Undergraduate Research

Software Engineering Laboratory, University of Science

Sept. 2023 – Present

- Engaged in a research-focused academic club under faculty mentorship.
- Studied Generative AI techniques such as Stable Diffusion and GANs.
- Participated in collaborative research and technical discussions.
- Contributed to scientific research projects in the field of Generative AI.

## Technical and Soft Skills

---

- **Back-End:** Python Django, Java, C++
- **Database:** Microsoft SQL Server, MongoDB, PostgreSQL
- **Other tools:** Overleaf
- **Strong presentation skills, Collaboration skills**

## Scholastic Achievements

---

- Awarded **Outstanding Student in Scientific Research Achievement** (2023)
- Honored in the **Top 100 Outstanding Students of Vietnam National University** (2023)
- Ranked **1st Best Student of Regular Program, Faculty of IT** (2021)
- Selected among **Top 5 Students** of the Faculty of IT (2021)
- Won **Third Prize** – Provincial Competition for Excellent Students in Informatics (2020)

## Key Academic Projects

---

### Embulk Input Marketo Plugin (Ruby, Embulk, Marketo REST API)

Role: Contributor — Open-source Developer

2024

- Contributed to an open-source Embulk plugin enabling data ingestion from Marketo using REST API.
- Improved support for bulk extraction of Leads, Activities, Campaigns, and Program Members with incremental ingestion.
- Collaborated on enhancing configuration parameters (batch size, polling interval, incremental filters) for scalability and reliability.
- GitHub: [embulk-input-marketo](#)

### Remote Control PC (Java, Socket Programming)

Team size: 6 — Role: Developer, Project Manager

May 2022

- Built an application to remotely control a PC (process listing, app listing, screenshots, shutdown/restart).
- Applied socket programming and Java toolkits in a team project for the Computer Networking course.
- GitHub: [Remote\\_Control](#)

### Chess Game (C/C++, SFML Library)

Team size: 5 — Role: Developer, Project Manager

Aug. 2022

- Developed a two-player chess game application for the Object-Oriented Programming course.
- Implemented game logic and user interface with SFML.
- GitHub: [Chess\\_OOP](#)

## Additional Certificates

---

- Completed **FPT AI Compass**
- Completed **AWS Academy Cloud Foundations**
- Completed **Intro to Machine Learning (Kaggle)**
- Completed **Unsupervised Learning, Recommenders, Reinforcement Learning (Coursera)**
- Completed **Advanced Learning Algorithms (Coursera)**
- Completed **Machine Learning Explainability (Kaggle)**
- Completed **Machine Learning Specialization (Coursera)**
- Completed **Ultimate AWS Certified AI Practitioner AIF-C01**
- Completed **AWS Certified Developer Associate**