Thi Thanh Hien Nguyen

06 95 10 70 82 — ntth2807vn@gmail.com — thi-thanh-hien-nguyen — neihhien.github.io

Profile — Recent graduate in Modeling and Artificial Intelligence with a robust background in statistics, machine learning, and deep learning. Passionate about advancing research through the application of mathematics and artificial intelligence in health and biology. Known for being curious, self-driven, adaptive, and highly collaborative.

Skills

- **Programming Languages:** Python, R, C, C++, SQL, Java
- **Data Science Tools:** Pandas, Numpy, Seaborn, Matplotlib, OpenCV
- **ML/DL Frameworks:** Scikit-learn, TensorFlow, PyTorch, HuggingFace
- **Techniques:** Deep Learning, Machine Learning, Computer Vision, Explainable AI, Optimization, Data Analysis
- **Languages:** French *(Full Professional Proficiency)*, English *(Autonomous)*, Vietnamese *(Native)*

Experience

Pierre Fabre Dermo-Cosmétique **Sep 2022 – Feb 2025**
Data Science Apprentice *Toulouse, France*
Focused on data-driven analysis, R&D, and applying artificial intelligence in the dermo-cosmetic industry.
Relevant missions:

Analysis of sensory profiles and consumer data:

- Reviewed literature to validate analysis methodologies.
- Improved and maintained an R Shiny application for statistical analysis.
- Developed a quality control tool for sensory test data.
- Built an application to monitor hairstylist performance.

Optisen Project - Predicting consumer preferences and skincare sensory profiles *(Engineering Thesis Project)*:

- Predicted consumers' overall appreciation based on sensory profiles.
- Developed a multi-objective model predicting sensory characteristics from chemical compositions.
- Integrated explainable AI methods (SHAP, LIME) for robust result interpretation.
- Designed a user interface to facilitate the new product formulation process.

Personal Projects

Independent Study: Multi-omics Data Analysis and Integration **Mar 2025**
Material from Summer School on Multi-omics Data Analysis and Integration, 2023

- Developed skills in analyzing and integrating RNAseq, metabolomics, lipidomics, and proteomics datasets.
- Applied statistical methods such as RGCCA and mixOmics to integrate and interpret complex data.
- Gained proficiency in network-based techniques such as WGCNA and SNF for biological insights.
- Contextualized data analysis results using metabolic models, semantic web tools, and knowledge graphs.

Exploration of Geometric Deep Learning **Feb 2024 – May 2024**

- Followed the AMMI (GDL100) course and tutorials on geometric deep learning.
- Developed a model to predict molecular properties using Graph Neural Networks (GNNs).

University Projects

Predicting molecular energy from their geometric structure **Apr 2024 – Jun 2024**

- Represented molecules geometrically using the Coulomb matrix and scattering.
- Trained a regression model under physical constraints to predict molecular energy.

Movie recommendation system based on posters and plots **Mar 2024 – May 2024**

- Developed image analysis (CNN) and text analysis models (NLP – Transformers).
- Developed an interactive interface using Gradio.

Education

INSA Toulouse & INP-ENSEEUHT **Sep 2022 – Feb 2025**
Double Degree in Modeling and Artificial Intelligence *Toulouse, France*

- **Relevant Courses:** Modeling and Scientific Computation, Statistics, Machine Learning, Deep Learning, Data Analysis, Technologies for AI (Recommendation System, NLP, Explainable AI)

References

Maugis-Rabusseau Cathy *Professor, INSA Toulouse, cathy.maugis@insa-toulouse.fr*
Deguilhen Perrine *User and Sensory Expert, Pierre Fabre Dermo-Cosmétique, perrine.deguilhen@pierre-fabre.com*