ISMAIL HATIM

Github | LinkedIn | ismail.hatim@lix.polytechnique.fr | +33 6 52 82 56 89

PhD Student at École polytechnique
Foundation Models for EHR (Graph Transformers, Diffusion, Multimodal Learning)

TRAINING & ACADEMIC HISTORY

École Polytechnique - PhD in Computer Science — **Palaiseau, France**
Thesis Title: A Graph Transformer Foundation Model for Sparse and Heterogeneous EHR Data — January 2026 - Today

- Development of a foundation model EHR data, in collaboration with LIX (Inria CEDAR) and the American Hospital of Paris.
- Design and benchmark of protocols across tasks (imputation, prediction) and institutions, with robustness and fairness evaluation.

CentraleSupélec - MSc in Artificial Intelligence — **Gif-sur-Yvette, France**
Research oriented Artificial Intelligence training in Machine Learning, Deep Learning and Stochastic Optimization — September 2024 - Present

- Core courses: Machine Learning, Decision Modeling, Optimization for Machine Learning and Deep Learning.
- Elective courses: Reinforcement Learning, Natural Language Processing and Network Science for Machine Learning.

ENSEA - École Nationale Supérieure De l'Électronique et de ses Applications — **Cergy, France**
Engineering training in Analog and Digital Electronics and Computer Science — September 2019 - February 2023

- Specialization in Computer Science and Systems.
- Courses taken: Digital Signal Processing, Software Engineering, Parallel Programming, Operating Systems and Algorithms.

LYCÉE CHAPTAL - Preparatory class for the Grandes Écoles — **Paris, France**
Maths Physics and Engineering Sciences then Mathematics Physics star class — September 2017 - June 2019

- Specialization in Computer Science.

PROFESSIONAL EXPERIENCE

INRIA - OPIS Team (CVN) in collaboration with Institut Imagine — **Gif-sur-Yvette, France**
Multimodal Graph Representation Learning & Missing Modalities Imputation — May 2025 - October 2025

- Implementation of a reproducible pipeline for multimodal EHR preprocessing, masking and patient-graph construction:
  - Design and implementation of a hybrid diffusion imputer.
  - Implementation of a conditional graph denoisers and a refinement loop.
  - Conversion of EEG, MRI and clinical annotations into graph features to integrate them into the pipeline.
  - Evaluation of imputation quality and downstream clinical performance with ablations and sensitivity studies to discretization.
- Extension of the fairness project on diffusion models on large graphs, quantifying the bias trade-off from graph structure and features:
  - Exploration of multiple diffusion models on large-scale graphs and on standard benchmark graphs.
  - Extension of evaluations to other downstream tasks (node classification and graph clustering).
- ML teaching support and practical lab supervision at a CentraleSupélec summer school.

POLARYS - Business Intelligence consulting firm — **Paris, France**
Business Intelligence Consultant affiliated with various missions — February 2023 - September 2024

- **Christian DIOR Couture: Tech Lead** within an AGILE project for data visualization:
  - Optimization of KPIs calculation times on PowerBI (DAX query execution times).
  - Integration of data into PowerBI data flows, and management of dataset updates for the Americas execution committee.
- **Eaton Corporation:** Configuration of payroll data flows using SaaS **beqom:**
  - Creation of the global architecture for data integration pipelines on Azure Data Factory.
  - Development of SQL stored procedures for ETL.
- **Ardian:** Migration of an integration solution (private equity company):
  - Integration architecture for the various flows (SSIS).
  - Parameterization of the solution using configuration files.

ACADEMIC AND TECHNICAL PROJECTS

Selected research projects demonstrating expertise in fairness, explainability and graph-based learning

CentraleSupélec - Lab Project: Fairness in Generative Graph Machine Learning — **Gif-sur-Yvette, France**

- Implementation of an evaluation pipeline of fairness in GeML models, focusing on architectures like VGAE and diffusion models.
- Integration of PyTorch Geometric to manage graph data structures and build GNN architectures for downstream tasks.

CentraleSupélec - Research Project: Explainability in AI Models — **Gif-sur-Yvette, France**

- Design of a framework to enhance interpretability in image classification by decomposing model decisions into spatial attributions.
- Application of the method on various vision models and validation of its robustness under common explainability methods.

SKILLS

Computer skills:

Programming Languages: Python, Java, SQL, C, C++

Libraries & Frameworks: PyTorch, PyTorch Geometric, TensorFlow, SciPy, Keras.

Languages: French & Arabic (Native), English (Advanced, TOEIC: 875), Spanish (Basic)