

# Khaoula BELAHSEN

linkedin.com/in/khaoula-belahsen/ | <https://github.com/khaoulabelahsen>

Email: khaoula.belahsen@proton.me

(+33) 617909049

## EDUCATION

---

### Ecole Normale Supérieure Paris-Saclay

Cachan, France

Masters in Mathematics, Vision, and Machine Learning. High Honors

2019 - 2020

Relevant coursework: Probabilistic Graphical Models, Statistical Learning, Advanced NLP for Text and Graph Data, Computer Vision

### ENSTA Paris / Institut Polytechnique de Paris

Palaiseau, France

BSc Applied Mathematics and Computer Science. GPA 3.82/4.3

2017 – 2020

MSc Statistics and Optimization, GPA 3.94/4.3

### Lycée Moulay Youssef

Rabat, Morocco

Intensive 2-year course (+80hours per week), GPA 3.99/4

2015 - 2017

Training in pure mathematics, physics, and computer science in preparation for competitive entrance exams to top French and Moroccan Engineering schools.

- Ranked 40/1779 nationally in CNC 2017 (Moroccan nationwide entrance exam)
- Top 10 Moroccan students for Mines-Ponts (French nationwide entrance exam)

## RESEARCH INTERESTS

---

- Single-cell multi-omics, cancer biology, immunology, protein and RNA binding, drug discovery process, graph neural networks, foundation models.

## HONORS & AWARDS

---

- 3rd prize at Citadel London Datathon 2020.
- French Government Excellence Scholarship awarded to the top 10 Moroccan students in French Schools (2017-2020)
- King Mohammed VI Foundation Scholarship awarded to the top 100 Moroccan high-school graduates (2015-2017)

## PROFESSIONAL AND RESEARCH EXPERIENCE

---

### AQEMIA (Biotech Startup - small molecule drug discovery)

Paris, France

*Machine Learning Engineer – Team Lead*

Nov 2022 – Present

- Applied Machine Learning methods to computational chemistry R&D problems (hit identification, ligand properties prediction..)
- Curated protein-ligand datasets and developed ML-based scorers to predict protein-ligand affinity (graph neural networks).
- Project Management: As a team lead, I supervised the R&D project road mapping, delivery, and documentation, as well as inter-team communication and presentations to C-level steering committees.

**EPIGENE LABS** (Precision oncology)

Paris, France

*Research Data Scientist*

March 2021 – Sept 2022

- Built data integration pipeline for harmonized clinical data and processed RNA-seq data.
- Trained ML models for patient survival prediction in collab with Institut Gustave Roussy.
- Cut by 75% time spent by biologists on clinical data QC through rules automation.
- Organized weekly paper reading group.

**KAYRROS** (Computer Vision)

Paris, France

*R&D intern*

May 2020 – Dec 2020

- Built a deep learning solution for automatic building detection on 2D/3D satellite image data.
- Implemented image resolution correction algorithm to improve quality of data inputs.

**Lady Davis Institute for Medical Research** (McGill University)

Montréal, QC

*Graduate Research Assistant to Dr. Celia Greenwood*

May 2019 – Aug 2019

- Translated biology requirements from physicians and biologists into statistical modeling.
- Modeled bias in DNA-methylation data based on GC counts.
- Improved a statistical interaction model for high features and low sample use cases (DNA methylation data) using custom regularizations and non-linear interaction terms.

**PRESENTATIONS**

---

- (Oral Presentation) **Belahsen, K.** (2023) Active Learning for Virtual Screening of Molecules in Drug Discovery. Paris Data Ladies Meetup 2023.

**TECHNICAL SKILLS**

---

- **Programming languages:** Python, R, SQL, some MATLAB
- **ML Engineering:** Pytorch and PyG, MLFlow, Pandas, PostgreSQL, Polars, Cloud Computing (AWS).

**COMMUNITY ACTIVITIES**

---

**Writer, Bits in Bio Newsletter***Since Feb 2024*

- Weekly curation and cover of the latest research papers and innovations in the Bio AI field
- Newsletter read by +7000 community members.

**Invited speaker, Boston Consulting Group, Womentoring 2023***Oct – Nov 2023*

- Mentoring program for girls in STEM engineering and business programs.
- Invited to speak about my experience and professional work.