KARIM BOUCHAARA

M2 BIOINFORMATICS & DATA SCIENCE STUDENT

*Looking for **6 month internship** (mid Jan 2026)*
M2 student with strong skills in Python, data analysis and deep learning, applied to real datasets (RNA folding Kaggle, climate modelling, omics data).
I now aim to join a data / AI team in a start-up to build and improve ML models and data pipelines.

CONTACTS

07 82 07 77 04

karimbchr4@gmail.com

71 rue Villeneuve, 92110 Clichy

https://github.com/karimbchr

https://karimbchr.github.io/

SKILLS

- **Programming**

Python, R, bash, SQL, HTML, Java

- **ML / Deep Learning**

Scikit-learn, PyTorch
BiLSTM, Conv1D, Transformer
Feature engineering, preprocessing, model evaluation

- **Data Analysis**

Data wrangling, workflows, pipelines

- **HPC**

Linux, SLURM, Singularity/Apptainer, GPU

HOBBIES

- Climbing (USMA Club)
- Solo travelling (Interail)
- Skateboarding

LANGUAGES

ENGLISH / PRO LEVEL

FRENCH / NATIVE

EXPERIENCE

Bioinformatics Intern

RPBS, PARIS

MARCH 2025 - JUNE 2025

- Implemented the RFpeptides deep-learning pipeline for cyclic peptide design
- Ran large-scale GPU jobs with SLURM / Apptainer
- Analysed structural scores (RMSD, pLDDT) and visualisations

Bioinformatics Intern

INRAE, SOPHIA-ANTIPOLIS

JUNE 2024 - AUGUST 2024

- Deployed tools for Iso-Seq PacBio data on an HPC platform
- Benchmarked parameters of minimap2 and gmap, visualised in IGV
- proposed improvements for long-read transcriptomics workflows

PROJECTS

RNA Folding prediction – Kaggle "Stanford Ribonanza"

- Built a Python pipeline to predict RNA chemical reactivity from sequence using BiLSTM / Transformer models
- Handled preprocessing (one-hot encoding, normalization, masking) and model evaluation

Climate-driven modelling of Ixodes scapularis

- Combined climate and tick-density data to model suitability across US counties; fitted Gaussian models and produced maps (R/ggplot2)

EDUCATION

Master in Bioinformatics – Biology IT

UNIVERSITÉ PARIS CITÉ – 2024–2026

Courses: bioinformatics, algorithms, peptide design, omics, analysis, statistics, ML

BSc in Life Sciences – Bioinformatics track

UNIVERSITÉ CÔTE D'AZUR (3RD IN CLASS) – 2021–2024

Courses: programming, genomics, molecular biology, statistics, modelling