

Massinissa Mouhoub

 Mikelenjilo |  Massinissa Mouhoub |  massinissa.mouhoub3@gmail.com |  +213 540 050 996 |   massi-dev.vercel.app |  Algiers, Algeria

RESEARCH SUMMARY

Machine learning researcher working on **recommender systems**. My M.Sc. thesis formulated therapeutic exercise prescription as **link prediction on a bipartite graph**, where a **graph attention network** outperformed a matrix factorisation baseline by **31% relative** on held-out data, and a learned feature-to-embedding encoder resolved the **cold-start** problem for unseen users. Co-author of a peer-reviewed IEEE conference paper. My research interests are graph-based and neural recommendation, cold-start and sparse-interaction settings, and evaluation methodology for ranking systems. I bring 3+ years of production software engineering — tested, documented, maintainable code — to research work.

EDUCATION

- M.Sc. Bioinformatics**, University of Science and Technology Houari Boumediene (USTHB), Algiers (2023–2025) — *thesis graded 17/20; ranked 2nd of 50 students*
- B.Sc. Computer Networking and Telecommunication Engineering**, USTHB, Algiers (2020–2023) — *final-year project graded 16/20; ranked in the top 5 of 51 students; 1st of 120+ in the first year*

PUBLICATIONS

- [1] K. Abbaci, F. M. Bouyakoub, I. Chegroune, **M. Mouhoub**. “A Smart System for Detecting High-Risk Areas of Epidemic Spread.” *12th IFIP International Conference on New Technologies, Mobility and Security (NTMS)*, Paris, France, 2025, pp. 292–299. IEEE. [doi:10.1109/NTMS65597.2025.11077002](https://doi.org/10.1109/NTMS65597.2025.11077002)
Bachelor's thesis project, developed jointly with I. Chegroune under the supervision of the two first authors. Built a reproducible Python pipeline over geospatial epidemiological data, applying unsupervised clustering (Scikit-learn) to identify areas most vulnerable to epidemic spread, with multi-algorithm benchmarking and statistical validation.

RESEARCH EXPERIENCE

- M.Sc. Thesis — Graph-Based Recommender System for Therapeutic Exercise Prescription** 2025 — 17/20
- **Supervisors:** Pr. Malika Mehdi (USTHB) and Pr. Khadidja Henni (Université TÉLUQ, Québec, Canada). Conducted in collaboration with a TÉLUQ laboratory and a team of rheumatologists and physiotherapists.
 - **Problem.** Formulated personalised exercise prescription for knee osteoarthritis as **link prediction over a bipartite patient–exercise graph** — 413 patients, 36 exercise categories (71 exercises), 9,376 observed interactions — built from 3D knee kinematic gait data captured by a KneeKG device.
 - **Model.** Implemented a **4-layer graph attention network (GAT)** in **PyTorch Geometric** with residual connections, batch normalisation, **focal loss** to handle interaction sparsity, negative sampling, and a hybrid MLP / dot-product / cosine decoder (132,545 parameters).
 - **Result.** Benchmarked against a **matrix factorisation** baseline on a held-out test set: **RPS 85.0% vs 64.7%**, a **31% relative improvement**. RPS is a composite metric I designed, 0.5 AUC + 0.3 AP + 0.2 Accuracy, weighted to prioritise threshold-independent ranking over raw accuracy under class imbalance.
 - **Cold start.** Designed an encoder network projecting a new patient's 70-dimensional clinical feature vector into the learned graph embedding space, enabling recommendations for users with no interaction history (validation RMSE 0.031).
 - **Ablations.** Isolated the contribution of individual architectural choices: residual connections (+7.4 RPS), MLP decoder (+4.3 RPS), and network depth (performance peaking at 4 layers, degrading beyond 5 through over-smoothing).
 - **Delivery.** Shipped as a working web application serving recommendations for both existing and previously unseen patients. Written in English; full text available on request.

ACADEMIC PROJECTS & TECHNICAL REPORTS

Unpublished coursework and project reports. Full texts (EN/FR) at massi-dev.vercel.app/papers.

Big Data Analysis with Apache Spark	Large-scale data processing and analysis using Apache Spark, covering distributed computing patterns and performance tuning. (2024)
Parallelising Smith–Waterman	Implementation and optimisation of the Smith–Waterman local sequence alignment algorithm using parallel computing techniques, with performance benchmarking. (2024)
Advanced Data Mining Techniques	Implementation and comparative analysis of data mining algorithms for pattern recognition, clustering, and predictive modelling across several datasets. (2024)
LCS via Genetic Algorithms	Approximate solution of the Longest Common Subsequence problem using genetic algorithms, with performance benchmarking against exact methods. (2024)
ML Fundamentals from Scratch	Neural network training loop with backpropagation, and linear regression with gradient descent, implemented from first principles in Python without ML libraries. [GitHub] [GitHub]

TECHNICAL SKILLS

Recommender Systems	Collaborative filtering, matrix factorisation, link prediction on bipartite graphs, cold-start handling, negative sampling, implicit-feedback modelling, ranking evaluation (AUC, average precision), composite metric design.
Machine Learning	Graph Neural Networks (GAT), attention mechanisms, deep learning, class-imbalance handling (focal loss), hyperparameter search, ablation studies, reproducible benchmarking, unsupervised clustering, statistical inference.
Algorithms & Computing	Parallel and distributed computing (Apache Spark, CUDA), genetic algorithms and metaheuristics, sequence alignment algorithms, complexity analysis and performance benchmarking.
Programming Languages	Python, Java, Kotlin, Dart, C, Go, JavaScript.
Libraries & Tools	PyTorch, PyTorch Geometric, Scikit-learn, Pandas, NumPy, Git , Linux/Bash, LaTeX , REST APIs, Django.
Software Engineering	Clean Architecture, Test-Driven Development (TDD), unit & integration testing, Agile/Scrum, technical documentation.

PROFESSIONAL EXPERIENCE

Mobile Engineer — YAY Global	July 2025 — present
– Diagnosed and resolved systematic production failures, reducing the crash rate from 19.8% to 1.05% by root-causing a native lifecycle defect — sustained analytical debugging against a real, noisy system at scale (60k+ users). – Architected a reusable component library adopted across the codebase, reducing feature delivery time and enforcing consistency — systems design under real maintenance constraints. – Collaborate within a multidisciplinary team of 10+ engineers, designers, and product staff, maintaining shared codebases under Git and delivering documented, production-quality software.	
Mobile Engineer — SlickPay, Yarbit, MIXIN (contracts)	March 2023 — May 2025
– Developed and maintained production software applying Clean Architecture and TDD with comprehensive unit and integration testing — the same discipline I apply to research code and reproducible experiments. – Authored a 30-page technical guide documenting system architecture, conventions, and reusable patterns, cutting onboarding time for new developers by half — written communication of complex systems to a mixed audience.	

LANGUAGES

French (fluent — TCF C2) | English (fluent) | Arabic (native) | Kabyle (native) | German (beginner)

REFERENCES

Available on request.