AmirHosein Ebrahimi

Ottawa, ON, Canada · amirh.ebrahimi1377@gmail.com
Google Scholar · linkedin.com/in/amirhosein-ebrahimi · github.com/EbrahimiAmirHosein · Portfolio

SUMMARY

AI/ML research engineer with an MASc in Biomedical Engineering, five years of research experience, and a publication record across computer vision, deep learning, computational pathology, omics, and neural systems. Experienced translating research questions into reproducible Python software, training and evaluating models with PyTorch and TensorFlow, debugging data and pipeline failures, and validating generalization on real-world external datasets.

TECHNICAL SKILLS

Machine Learning: PyTorch, TensorFlow, Keras, Scikit-learn, XGBoost, Hugging Face Transformers; deep learning, supervised and unsupervised learning, classification, regression, clustering, representation learning and weak supervision

Computer Vision & Generative AI: CNNs, U-Net, object detection, image segmentation, computational pathology, whole-slide imaging, self-supervised learning, domain adaptation, foundation-model representations, GANs and diffusion-model concepts

Validation & Statistics: Experimental design, benchmark models, cross-validation, hypothesis testing, ablation studies, calibration concepts, explainability, subgroup analysis, error analysis, external validation and robustness testing

Programming & Data: Python, SQL, R, C++, Java, Bash/Linux, Pandas, NumPy, SciPy, PostgreSQL/PostGIS, relational data, data cleaning, feature engineering, Git and object-oriented programming

Research & Reporting: Research-paper implementation, reproducible experimentation, model documentation, technical reports, statistical visualization, presentations and communication with multidisciplinary stakeholders

Systems & Delivery: Docker, FastAPI, REST APIs, GPU-based training, batched data loading, cached representations, AWS and Azure concepts, MLflow, CI/CD, Databricks and distributed-computing concepts

EXPERIENCE

Mer Lab, University of Ottawa **2023 – Present**
Research Assistant (MASc thesis student, 2024 – 2026) *Ottawa, Canada*

- Design and implement machine-learning methods across clinical, genomic, transcriptomic, and imaging datasets using Python, PyTorch, TensorFlow, Scikit-learn, XGBoost, Pandas, and NumPy.
- Develop validation methodologies using benchmark models, controlled experiments, cross-validation, ablations, external cohorts, subgroup inspection, explainability methods, and systematic error analysis.
- Investigate model generalization across datasets and institutions, identify distribution and acquisition shifts, evaluate mitigation strategies, and document remaining limitations.
- Built OmicSieve, an unsupervised pipeline that transforms RNA-seq profiles of approximately 19,000 genes into compact representations for downstream analysis [code: OmicSieve].
- Contributed to drug-response modeling integrating CCLE data from 1,145 cell lines, GDSC data from 1,084 cell lines, and responses across 38 drugs; supported feature selection, statistical analysis, benchmarking, and reporting.
- Communicate methods, findings, risks, and recommendations through technical documentation, presentations, publications, and collaboration with computational and biomedical researchers.

Huma **2023 – 2024**
Solution Specialist *London, United Kingdom*

- Investigated and debugged software and data issues affecting remote patient-monitoring and clinical-research products in a regulated, privacy-sensitive production environment.
- Worked with engineering and product teams to reproduce issues, clarify requirements, document findings, surface risks, and communicate reliable resolutions to technical and non-technical stakeholders.

Institute for Research in Fundamental Sciences (IPM) **2019 – 2024**
Research Assistant *Tehran, Iran*

- Implemented and evaluated numerical and deep-learning algorithms in Python and C++, systematically comparing predictive and computational tradeoffs through controlled experiments and benchmarking.

SELECTED PROJECTS

LungInsight [Website] · *Python, FastAPI, Docker, Model Validation*

- Designed and built a Dockerized annotation application and connected expert annotations to data preparation, prediction, validation, and externally sourced test data.
- Evaluated model behavior on data from a different laboratory, investigated acquisition-related performance shifts, and used the findings to improve data checks and robustness analysis.

LIMIT: Tile-Efficient Weakly Supervised Learning for Molecular Prediction · *Python, PyTorch, Large-Scale Data*

- Built a reproducible workflow over approximately four million image tiles, using batched loading and cached foundation-model representations to support efficient experiments across patient cohorts.
- Designed a self-supervised selector, compared against benchmark methods, and evaluated cross-domain generalization and domain-adversarial learning. Code is unavailable while the associated paper is under review.

RAG Workflow [Website] · *Python, Hugging Face, Transformer Embeddings*

- Built a retrieval-augmented generation workflow using patient history to contextualize prescription guidance and ground a patient-facing chatbot in retrieved information.

EDUCATION

MASc, Biomedical Engineering (Specialization in Artificial Intelligence) **2024 – 2026**
University of Ottawa *Ottawa, Canada*

- GPA 4.0/4.0. Thesis: *Tile-Efficient Weakly Supervised Learning for Molecular Prediction from Histopathology*.

BSc, Computer Engineering **2017 – 2022**
Shahid Beheshti University *Tehran, Iran*

PUBLICATIONS

* First author.

1. * High-performance and configurable SW/HW co-design of biologically plausible neural networks. *ACM Transactions on Reconfigurable Technology and Systems (TRETS)*, 2026 (accepted).
2. * OmicSieve: A general omics framework for compact component embeddings. *IEEE CIBCB*, 2026 (in press).
3. Artificial intelligence-enabled histological analysis in pre-clinical respiratory disease models: A scoping review. *European Respiratory Review*, 35(180), 250243, 2026.
4. * From data annotation to AI prediction: Streamlining histopathology analysis in acute respiratory distress syndrome. *38th Canadian Conference on Artificial Intelligence*, 2025.
5. * Basics of machine learning. In *Dimensionality Reduction in Machine Learning*, pp. 3–38. Elsevier, 2025.
6. Deciphering genomic correlates of differential treatment response kinetics in multiple myeloma. *Clinical Lymphoma, Myeloma & Leukemia*, 25(Suppl. 2), S153–S154, 2025 (conference abstract).
7. Text-mining-based feature selection for anticancer drug response prediction. *Bioinformatics Advances*, 4(1), vbae047, 2024.
8. * HA-BSN: Hardware acceleration of bio-SFA and bio-NICA, biological neural networks, on FPGA with HLS. *SSRN*, 2023 (preprint).