Bahar Emami Afshar

Machine Learning Engineer | MSc AI | Published Author | Inventor on Filed Intellectual Property

171 O'Conner St, Ottawa, Ontario | *613-447-1656* | *Google Scholar* | *baharafshar2079@gmail.com* | *Linkedin* | *Github* | *Website*

Professional Summary

Research-focused ML scientist with a strong software development background and a passion for solving real-world problems at scale. Published 3+ peer-reviewed papers on anomaly detection, active/semi-supervised learning, and explainability in financial applications. Proven ability to build scalable ML pipelines and apply deep learning to large, imbalanced datasets. Experienced with distributed systems, Docker/Kubernetes, and cloud-based ML workflows. Skilled in Python, C/C++, and object-oriented design. Open to relocation in Ontario.

Technical Skills

Languages : Python, C/C++, Java, Javascript, R, SQL, Verilog, HTML, XML, JSON, CSS.
Web Frameworks : React Native, Node.js, Flask, Django, Spring, Bootstrap.
DevOps : Git, Docker, Kubernetes, Kafka.
Databases : MySql, HibernateORM, Elastic Search, MongoDB, Neo4j.
ML Frameworks : Pandas, NumPy, TensorFlow, Transformers, PyTorch, Scikit-Learn, Hugging Face.
ML Expertise :Anomaly Detection, Natural Language Processing, Deep Learning, Active Learning, Online Learning, Semi-Supervised Learning, Large Language Models, Data Analysis, Big Data, Explainable AI, Association Rule Mining, and Meta Learning.

Experience

ML Engineer Researcher (MITACS Internship) | *H3M Analytics Inc.*|Ottawa, ON *09/2023 – 08/2025*

- Supervisor: Dr. Paula Branco (University of Ottawa)
- Developed **X-ITERADE**, an explainable unsupervised anomaly detection framework with a modular backend for fraud detection that identifies high-quality suspicious cases without requiring labeled data, and achieved a **15-times improvement** in the imbalance ratio while maintaining flexibility under labeling budget constraints.
- Created ALISA, an innovative iterative learning pipeline combining active learning, semi-supervised learning, and data augmentation with dynamic weighting and feedback loops, boosting **F1-score by 22%** in highly imbalanced one-class scenarios.
- Developed an explainability module for group behavior analysis in financial datasets, leveraging **LLaMA**, **Mistral**, and **GPT** models with **in-context learning** and **prompt tuning**, achieving **90% accuracy** in interpreting fraud patterns.
- Designed and implemented ML models for extreme class imbalance (0.077%) using advanced algorithms including XGBoost, LightGBM, autoencoders, and Transformer-based architectures tailored for anomaly detection and learning from limited labels.
- Built end-to-end deep learning pipelines in Python using Pandas, Scikit-learn, Hugging Face, and PyTorch, ensuring reproducibility and robustness through version-controlled codebases (Git) and thorough documentation.

ML Engineer Intern | *Peppy Digger* | Tehran, Iran *05/2021 – 09/2021*

- Developed a **3-class sentiment analysis** model on Persian Twitter data using FastText embeddings, TF-IDF features, and Bidirectional GRU networks to capture rich language context.
- Applied class imbalance handling techniques, including dynamic class weighting and resampling, boosting **accuracy by 10%** while ensuring balanced performance across classes.
- Built and trained deep learning pipelines in Python with **TensorFlow/Keras**, optimizing with **Adam** and tracking metrics like accuracy and AUC.

Education

University of Ottawa — **M.Sc. in Computer Science - AI Applied Concentration** | GPA : A+ *09/2023 – 08/2025*
University of Tehran — **B.Sc. in Computer Engineering** | Last two years GPA : A+ *09/2018 – 07/2023*

Projects

Semi-Supervised Learning for Bank Marketing | *Machine Learning, Python, Numpy, XGBoost, LGBM* *03/2024*

- Compared and implemented multiple semi-supervised algorithms using models such as GBT, SVC, KNN, and MLP, improving accuracy from 76% to 88% on imbalanced marketing data.

Online Learning for Intrusion Detection | *AI for Cybersecurity, Python, Numpy, Apache Kafka* *11/2023*

- Developed a real-time intrusion detection system with Apache Kafka and online learning algorithms using XGBoost, achieving 87% F1-score while adapting to data drift and evolving threats.

COVID Detection from Lung Images | *Python, Pytorch, Tensorflow* *11/2020*

- Built and benchmarked CNNs in NumPy, PyTorch, and TensorFlow for chest X-ray classification, reaching 94% accuracy using data augmentation and tailored loss functions.

IMDB Clone (IEMDB) | *Internet Engineering, Java, Spring Boot, React, Docker, Kubernetes* *05/2022*

- Designed and deployed a full-stack movie platform with Spring Boot, React, Docker, and Kubernetes, including a collaborative filtering-based recommender system.

Publications & IP

UOttawa & H3M Analytics Inc. *2024-2025*

- Published 3+ peer-reviewed papers on fraud and malware detection using explainable and label-efficient AI systems; submitted one patent on a two-stage anomaly detection and explanation framework. Full list available on Google Scholar.