+46767194735
Stockholm, Sweden
ali.bm74@gmail.com

# Ali Beikmohammadi

AI Tech Lead | Applied AI Systems

linkedin.com/in/alibm95
github.com/AliBeikmohammadi
Personal Website, Google Scholar

AI Tech Lead and formally trained ML Engineer with 7+ years of experience designing, deploying, and operationalizing production-grade machine learning and generative AI systems. Proven track record of leading cross-functional initiatives and collaborating with industry and research partners to deliver scalable AI solutions across industrial, product, and enterprise domains. Strong focus on ML platforms, MLOps, and distributed systems, with hands-on experience spanning computer vision, agentic AI, and large-scale LLM pipelines. Known for translating ambiguous business problems into robust technical solutions, driving alignment across stakeholders, and enabling teams through clear communication, ownership, and sound engineering practices.

## Technical Skills

| | |
|---|---|
| **Technical Leadership:** | Solution design, Technical standards, Code review, Technical debt management |
| **Governance:** | Responsible AI, Risk assessment, Self-service enablement, Cost optimisation |
| **Programming Languages & Scripting:** | Python, CUDA, OpenMP, MPI, MATLAB, Bash/Shell Scripting |
| **Machine Learning & AI Frameworks:** | PyTorch, TensorFlow, Keras, Scikit-learn, Ray Tune |
| **MLOps Tools & Frameworks:** | TensorFlow Serving, MLflow, Flyte |
| **Big Data, Databases & Distributed Systems:** | SQL, MySQL, PostgreSQL, MongoDB, Kafka, Apache Spark, Hadoop, Hive |
| **Data Analysis, Processing & Visualization:** | Pandas, Seaborn, Plotly, Matplotlib, Power BI |
| **Cloud Platforms & HPC Systems:** | Microsoft Azure, Databricks, Copilot Studio, Power Automate, HPC clusters |
| **DevOps, CI/CD & Orchestration Tools:** | Docker, Kubernetes, Git, GitLab CI/CD, Azure DevOps |

## Professional Experience

### AI Tech Lead — Sep. 2026 — Present

***Sandvik/Seco Tools AB — AI & Automation Center for Enablement*** — ***Stockholm, Sweden***

- **Technical ownership and solution quality**
  - Own day-to-day solution design and technical decision-making across the team, aligning delivery with architectural direction, engineering standards, security requirements and governance.
  - Define the technical quality criteria — testing, performance, security, maintainability, scalability, cost and operational readiness — that solutions are held to.
  - Review complex and high-risk solutions before production release, escalating where specialist architectural, security, privacy or legal review is needed.
  - Promote reuse of approved datasets, components, services and patterns over duplicated one-off builds.
- **AI governance and Responsible AI**
  - Define practical guardrails for safe, scalable development across AI, GenAI, analytics, automation and low-code tooling.
  - Ensure AI solutions meet requirements for Responsible AI, security, privacy, data protection, human oversight and regulatory compliance.
  - Establish evaluation, testing, monitoring and lifecycle-management practice for models, prompts, agents and related components.
  - Balance self-service flexibility against the controls that keep solutions supportable and compliant.
- **Operational excellence and lifecycle management**
  - Own the team's approach to production stability, incident response, technical debt and continuous improvement.
  - Ensure every production solution has clear ownership, documentation, monitoring, support arrangements and a lifecycle plan.
  - Drive root-cause investigation of significant or recurring incidents through to resolved improvement actions.
  - Support cost monitoring and optimisation across AI, automation, data and cloud workloads.
- **Technical roadmap and risk**
  - Contribute to the AI and automation technical roadmap by identifying capability gaps, platform needs and modernisation opportunities.
  - Identify and communicate technical risk, dependencies and architectural constraints, with recommended mitigations.
  - Give Product Owners the technical consequences of prioritisation decisions without taking ownership of the backlog.
  - Feed implementation experience back into the evolution of architectural standards and platform capability.
- **Team enablement and ways of working**
  - Coach engineers, analysts and citizen developers through design reviews, code reviews and knowledge-sharing.
  - Distribute technical ownership deliberately to reduce dependency on individual specialists.
  - Build capability through documented patterns, technical demonstrations and structured knowledge transfer.

- Provide technical leadership without relying on formal authority.

- **Academic collaboration**
  - Guide collaboration with universities, researchers and Master's thesis students — shaping suitable AI topics, providing technical supervision and evaluating outcomes.
  - Connect promising academic work to real business challenges, technical roadmaps and experimentation inside the organisation.
  - Ensure academic collaboration meets requirements for confidentiality, intellectual property, information security and Responsible AI.

- **Platform and MLOps Ownership**
  - Drove adoption of scalable MLOps patterns across Azure ML and Databricks, including MLflow-based experimentation, model registry practices, Vector Search, RAG architectures, and Unity Catalog.
  - Authored internal standards and guidelines for CI/CD for ML, model lifecycle management, and deployment readiness.
  - Collaborated with DevX and platform teams to improve reliability, monitoring, reproducibility, and governance of production ML systems.

## Senior AI Engineer — Jun. 2025 — Oct. 2026

***Sandvik/Seco Tools AB — AI & Automation Center for Enablement*** — ***Stockholm, Sweden***

- **AI Enablement, Strategy, and Cross-Org Delivery**
  - Core contributor to the AI Center for Enablement, acting as an internal AI consultant, defining how AI is scoped, delivered, and operationalized across HR, L&D, Operations, Finance, Global Supply Chain, IT, Sales Operations, R&D, and Logistics.
  - Led AI initiative intake and use case discovery, shaping problem definitions, feasibility assessments, roadmaps, and execution plans across multiple domains.
  - Acted as a bridge between business leadership, product owners, and platform teams to align AI initiatives with strategic objectives and data readiness.

- **Agentic AI, Conversational Systems, and Automation**
  - Led design and delivery of multiple conversational and agent-based AI systems using Microsoft Copilot Studio, Databricks, and Azure, serving customer-facing, sales, operational, and HR use cases.
  - Developed and operationalized domain-specific AI agents, including:
    - **Customer-facing product discovery** — public website chatbot connected to marketing PDFs, product databases, and web content for product discovery and customer support.
    - **Manufacturing diagnostics** — internal chatbot for diagnosing machine and process errors in manufacturing environments.
    - **Sales enablement** — sales enablement chatbot providing facts, references, and success stories to support outreach to new customers.
    - **Learning and development** — HR and L&D chatbot analyzing employee profiles and recommending personalized learning paths aligned with the 70/20/10 learning model.
  - Integrated AI agents with Power Automate, evaluated trade-offs between traditional RPA and agentic automation, and documented licensing, governance, and operational best practices.

- **Product Intelligence, Generative AI, and Content Automation**
  - Led an AI-driven product description initiative to generate short and long product descriptions for 35K+ products using LLM-based pipelines.
  - Designed and implemented scalable data and generation workflows in Databricks and Azure, integrating structured product data and unstructured marketing content.
  - Defined extensive business rules, quality guidelines, and validation logic to ensure brand consistency, regulatory compliance, and high-quality outputs.
  - Collaborated with product, marketing, and IT teams to operationalize live content generation and continuous updates.

- **Industrial AI Systems (Automated Tool Wear Inspection)**
  - Technical lead, owning the end-to-end computer vision pipeline from data annotation and dataset strategy to model development and cloud deployment.
  - Designed and deployed object detection models (YOLO, Faster R-CNN) for industrial defect detection, while establishing annotation standards, QA processes, and dataset versioning.
  - Drove large-scale data annotation efforts by defining labeling guidelines, running annotation workshops, coordinating quality assurance, and establishing dataset versioning practices.
  - Contributed to architecture blueprint, risk management, functional and non-functional requirements, and migration of AI workflows from on-prem to Azure.

## ML Research Engineer
**Sep. 2021 — May 2025**

***Stockholm University — Data Science Research Group*** ***Stockholm, Sweden***

- **Reliable Adaptive Predictive Maintenance and Intelligent Decision Support Systems**
  ***Collaboration with SCANIA CV AB and Linköping University***
  - Collaborated with cross-disciplinary teams, including Applied Scientists, Software Engineers, and Analysts, to deliver ML solutions that directly impacted business KPIs.
  - Led the development of a library, proposed and developed MM K-means, a new real-time inference clustering method capable of handling up to 50% missing data while maintaining comparable accuracy to full-data K-means.
  - Designed and developed a novel cost-sensitive Transformer model, cutting truck braking system costs by 15%.
  - Oversaw implementations with a team of engineers and research scientists.
- **Mitigating Toxicity and Bias in Large Language Models via Reinforcement Learning**
  ***Collaboration with SINTEF AS and KTH University***
  - Developed a model-agnostic RL framework to minimize toxicity in LLMs across architectures like LLAMA, GPT, BERT, and RoBERTa.
  - Evaluated performance using diverse toxicity metrics, achieving significant reductions in Toxicity, Profanity, Sexually Explicit Content, Flirtation, Threats, Insults.
  - Conducted experiments on fairness and bias mitigation, demonstrating improved inclusivity and fairness scores in generated outputs.
- **AI-based Control and Coordination for Smart Converters**
  ***Collaboration with Hitachi Energy and KTH University***
  - Developed and implemented a multi-agent deep RL (MADDPG) system for smart converters with Hitachi Energy.
  - Enhanced Actor-Critic RL models by introducing Accelerated-TD3, achieving a 200% improvement in learning efficiency by integrating prior knowledge.
  - Pioneered a communication-efficient federated learning framework for MARL, improving collaboration efficiency by 40x.
  - Investigated the finite-sample convergence of federated learning algorithms like Error-feedback FedProx and FedAvg, achieving 80% faster convergence without data similarity, enhancing federated learning scalability.
  - Examined the non-asymptotic convergence of Distributed Momentum Methods under biased gradient estimations from compression/clipping.
- **Deep Reinforcement Learning for Sustainable Power System**
  ***Collaboration with KTH University and the University of California***
  - Created a PPO-based deep RL model for sustainable power systems, reducing energy consumption by 16%.
  - Designed a new environment based on Gymnasium for evaluating multi-agent RL algorithms, improving development cycles.
  - Introduced TA-Explore, a novel exploration strategy in RL, accelerating convergence rates in simulation environments by 150%.
- **Supervised 30+ Master's students in AI and Health Informatics** at Stockholm University and Karolinska Institutet, applying advanced ML techniques (deep learning, Transformers, multimodal and reinforcement learning) to diverse domains including healthcare, forecasting, recommendation systems, network optimization, planning, agriculture, and fairness.
- **Course Instruction - Current Research and Trends in Health Informatics**, with a focus on time-series analysis, deep learning methodologies, and reinforcement learning techniques. Instructed 91 students.

## Data Scientist
**Sep. 2019 — Oct. 2021**

***MCI — Iran's largest telecommunications operator with 60M+ subscribers and a leader in digital services.*** ***Tehran, Iran***

- **Fraud Detection and Risk Assessment**
  - Worked as the lead for a ML team of six, collaborating with feature teams for over six months. Worked closely with product, design, data scientists, and engineering managers in scoping initiatives using ML as a solution, including creating problem definition specifications and managing stakeholder relationships.
  - Developed pipelines for data collection, preprocessing, feature engineering, training, evaluation, deployment, batch predictions/inference of advanced ML models for fraud detection, resulting in a 40% reduction in false positives and saving $1M in potential losses while also streamlining processes and reducing implementation time by 30%.
  - Created a zero to one specification on comprehensive risk assessments across multiple business entities, leading to more accurate and actionable insights.
  - Developed software prototypes and demonstrations with a team, leveraging MLOps principles—including CI/CD pipelines, automated workflows, and containerization (Kubernetes and Docker)—to streamline the deployment, monitoring, maintenance, and troubleshooting of ML models in production at scale.
- **Customer Churn Prediction**

- Contributed to content, teaching, and tutoring for internal data and ML training programs for engineers. Additionally, led reading groups focused on ML topics within various teams and product areas.
- Led initiatives to identify performance gaps, co-designed and rolled out a system implementing SOTA ML models (CatBoost, LightGBM, XGBoost, Random Forest) to predict Telecom customer churn, improving prediction accuracy by 12%.
- Addressed data imbalance issues using SMOTE and ADASYN, resulting in more robust model performance on minority classes.

- **Supply Chain Optimization**
  - Onboarded new team members ranging from junior to staff level and mentored engineers on leading initiatives, including problem scoping, test specification development, and defining deliverables.
  - Scaled ML deployments to handle large-scale data processing and transactions. Troubleshot deployment issues, ensuring 99.999% uptime and performance consistency.
  - Built and serviced an optimized order lead time prediction system in supply chain management using Linear Regression, Decision Tree Regression, and Gradient Boosting Regression models, reducing lead time variance by 25% and improving operational efficiency.

## Projects

- Engineered a personalized treatment recommendation system using Decision Tree, Random Forest, and XGBoost, increasing treatment efficacy for obstructive sleep apnea patients to 87%.
- Developed a metaheuristic algorithm that optimized energy consumption in cloud data centers, reducing energy usage by 27% while maintaining performance.
- Created a hierarchical RL algorithm for network routing, boosting throughput by 29%, largely enhancing network efficiency.
- Optimized resource allocation with deep RL (DQN), increasing network capacity by 34%, enhancing communication efficiency.
- Developed a deep RL framework (SAC) tailored for vehicle fuel economy, reducing fuel consumption by 15%, contributing to more sustainable transportation solutions.
- Engineered a MaxViT-based Vision Transformer for leaf disease detection, achieving a 98.25% accuracy rate.
- Led the development of an Automatic Fused Multimodal DL model, boosting plant identification accuracy by 11.07%.
- Engineered an RNN model combining GRU and LSTM architectures, increasing air pollution forecasting accuracy by 8%, aiding in more timely public health responses.
- Created a model using a Time-aware Latent Hierarchical approach, improving housing price prediction accuracy by 18%, leading to more reliable market forecasts.

## Open Source Software Contributions

- **Per-FedAvg-PG** (Creator) Implementation of personalized federated reinforcement learning via model-agnostic meta-learning.
- **Accelerated-Actor-Critic** (Creator) Implementation of accelerated SAC/TD3 via pseudo-labels derived from prior knowledge.
- **TA-Explore** (Creator) Implementation of human-inspired framework to accelerate reinforcement learning.
- **OpenNARS-for-Applications** (Contributor) Implementation of a non-axiomatic reasoning system.
- **FedAlgo_WO_DataSim** (Creator) Implementation of FedAvg, error-feedback FedAvg, FedProx, and error-feedback FedProx.
- **DistributedSGDM** (Creator) Implementation of distributed momentum methods under Top-k and clipped sparsification.
- **CompFedR** (Creator) Implementation of compressed federated reinforcement learning with a generative model.

## Education

### Ph.D. in Computer and Systems Sciences — Sep. 2021 — Oct. 2026
***Department of Computer and Systems Sciences, Stockholm University*** — ***Stockholm, Sweden***

- **Relevant Courses:** Reinforcement Learning, Artificial General Intelligence, Majorize-Minimization Optimization with ML Applications, High Performance Computing

### M.Sc. in Electrical Engineering Digital Electronic Systems — Sep. 2017 — Sep. 2019
***Department of Electrical Engineering, Amirkabir University of Technology*** — ***Tehran, Iran***

- **Relevant Courses:** Machine Learning, Neural Networks, Deep Learning, Statistical Pattern Recognition, Machine Vision

### B.Sc. in Electrical Engineering Electronics — Sep. 2013 — Jul. 2017
***Department of Electrical Engineering, Bu-Ali Sina University*** — ***Hamedan, Iran***

- **Relevant Courses:** Calculus, Algebra, Statistics & Probability

## Selected Publications

1. **Ali Beikmohammadi**. "Toward Scalable, Robust, and Efficient Reinforcement Learning: Theoretical Foundations and Algorithms for Sample-Efficient, Communication-Aware, and Heterogeneity-Robust RL." Stockholm University, 2026. [**Doctoral thesis**]
2. **Ali Beikmohammadi**, Mohammad Hosein Hamian, Neda Khoeyniha, Tony Lindgren, Olof Steinert, and Sindri Magnússon. "A Cost-Sensitive Transformer Model for Prognostics Under Highly Imbalanced Industrial Data." Cluster Computing. 2026.
3. **Ali Beikmohammadi**, Sarit Khirirat, and Sindri Magnússon. "Distributed Momentum Methods Under Biased Gradient Estimations." IEEE Transactions on Control of Network Systems. 2025.
4. **Ali Beikmohammadi**, Sarit Khirirat, Peter Richtárik, and Sindri Magnússon. "Collaborative Value Function Estimation Under Model Mismatch: A Federated Temporal Difference Analysis." European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECML-PKDD). 2025. [**Selected in the top 24% of submissions**]
5. **Ali Beikmohammadi**, and Sindri Magnússon. "Human-inspired framework to accelerate reinforcement learning." The Journal of Supercomputing. 2025.
6. **Ali Beikmohammadi**, Sarit Khirirat, and Sindri Magnússon. "Compressed Federated Reinforcement Learning with a Generative Model." European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECML-PKDD). 2024. [**Selected in the top 24% of submissions**]
7. **Ali Beikmohammadi**, Sarit Khirirat, and Sindri Magnússon. "On the Convergence of Federated Learning Algorithms Without Data Similarity." IEEE Transactions on Big Data. 2024.
8. **Ali Beikmohammadi**, and Sindri Magnússon. "Accelerating Actor-Critic-Based Algorithms via Pseudo-Labels Derived From Prior Knowledge." Information Sciences. 2024.
9. **Ali Beikmohammadi**, and Sindri Magnússon. "Ta-Explore: Teacher-Assisted Exploration for Facilitating Fast Reinforcement Learning." Proceedings of the 2023 International Conference on Autonomous Agents and Multiagent Systems (AAMAS). 2023.
10. **Ali Beikmohammadi**, Karim Faez, and Ali Motallebi. "Swp-Leafnet: A Novel Multistage Approach for Plant Leaf Identification Based on Deep CNN." Expert Systems with Applications. 2022.