

Moeen Bagheri

AI Engineer | Toronto, ON, Canada

moinbag@gmail.com | moeenbagheri.com | github.com/bagherig | linkedin.com/in/moeenbagheri/

SUMMARY

AI Engineer with 12+ years of experience and a Master of Science in Data Science and Analytics, building production LLM agents. Technical lead for FOX, BASF's AI assistant serving 4000+ users and winner of BASF's 2025 Digital Award. Expertise in LangGraph agent and tool design, evaluation, LLM output grounding, and agentic coding.

EXPERIENCE

Senior AI Engineer

Oct 2024 - Present

BASF – Toronto, ON, Canada

Lead Engineer for FOX Talk2Data, BASF's flagship AI data assistant used by 4000+ users. Winner of BASF's 2025 Digital Award.

- **Enterprise AI Assistant:** Technical lead for design and development of **FOX**, BASF's AI assistant serving **4000+** users. Winner of BASF's 2025 **Digital Award**.
- **Natural Language to SQL:** Designed and developed a LangGraph agent that translates natural-language questions into Databricks SQL queries and returns tables and charts, runs deterministic analysis of the data, and answers from unstructured reports using RAG. New data and RAG sources onboard through configuration alone.
- **Grounded Executive Briefing:** Shipped a monthly AI briefing for finance leadership (live since Sep 2026): deterministic SQL finds which KPIs moved and why, the LLM writes only the narrative, and validation rejects unsupported figures. Built on a shared analysis engine that also powers the chat assistant, so both always report the same figures.
- **Dashboard-Aware Chat:** Connected the assistant to the *Power BI* report a user has open, passing the page's metadata from a *React* frontend through a *FastAPI* backend to the agent, so answers reflect the data on screen.
- **Evaluation:** Built an evaluation framework with automated tests that track SQL correctness, tool selection, and latency for each assistant, catching regressions before release.
- **Agentic Coding System:** Built an agentic coding assistant on top of GitHub Copilot and Cursor that autonomously implements features, runs tests, creates pull requests, and reviews code across production repositories. Architected with specialized sub-agents, instructions, skills, and custom prompts, enabling a single engineer to operate with the throughput of a small team.
- **AI Delivery Leadership:** Co-owner of the Forward Deployed Engineering and Delivery Acceleration work package in BASF, which aims to accelerate the delivery of AI-powered features to production.

Generative AI Engineer

Sep 2023 - Oct 2024

YT Copilot – Toronto, ON, Canada

Developed a Chrome extension using Generative AI to enhance YouTube experience.

- **YouTube LLM Agent:** Developed a Chrome extension that skips promotional segments, summarizes videos, and enables semantic search over YouTube content using an agentic LLM.
- **Context Window Management:** Used a *MapReduce*-style chunking and aggregation pipeline to efficiently handle LLM context limitations for long YouTube videos.
- **Fullstack Development:** Built a *JavaScript*-based frontend integrated with YouTube's UI and deployed a *Flask* / *PostgreSQL* backend on *Azure Kubernetes Service (AKS)* with *Docker*, plus *Ory Kratos* authentication and *Stripe* billing.
- **Finetuning & Optimization:** Improved response quality and reliability using tool calling, structured outputs, fine-tuning, and prompt engineering techniques such as chain of thought prompting.
- **CI/CD Pipeline:** Implemented a *CI/CD* pipeline with *Azure CLI* for automated testing, deployment, and monitoring.

Machine Learning Engineer

Oct 2021 - Jul 2023

Super GeoAI Technology Inc. – Toronto, ON, Canada

Led team of 7+ engineers developing computer vision models for agriculture.

- **Computer Vision:** Led 7+ engineers to develop object detection models using *PyTorch/TensorFlow* for seed grading and crop yield estimation from phone/drone images, improving efficiency for agricultural partners by over **80%**.
- **Model Enhancement:** Applied transfer learning (*COCO* and *Pascal VOC*) and performed hyperparameter optimization (*Ray Tune*), improving accuracy by **10%** for seed grading and **15%** for wheat head detection.
- **System Design:** Deployed using *Flask* integrated with a *NextCloud*-based storage platform and *Ory Kratos* authentication.
- **Leadership:** Conducted technical interviews, led team meetings, and enforced high-quality implementation standards through structured code reviews.

Machine Learning Engineer

Jul 2021 - Jan 2022

IBM – Toronto, ON, Canada

Used NLP to identify conflicting or duplicated entries in SRS documents.

- **NLP:** Used *BERT*-based models to perform *NLP* on Software Requirements Specification (SRS) documents to detect conflicting and duplicate requirements.
- **Model Optimization:** Achieved a **20%** increase in F1-score through fine-tuning and hyperparameter optimization of a Microsoft pre-trained model.

Instructor - Practical Deep Learning (CIND850)

Sep 2021 - Dec 2021

Ryerson University – Toronto, ON, Canada

Mentored 20+ graduate students on applied deep learning projects.

- **Mentorship:** Mentored 20+ graduate students on applied deep learning projects, reviewing model architectures and ensuring best practices.
- **Assessment:** Graded assignments and exams and provided targeted feedback on deep learning implementations and results.
- **Curriculum Development:** Designed lab exercises/solutions, emphasizing practical applications of deep learning.
- **Lecture Delivery:** Delivered lectures and tutorials on core deep learning concepts, ensuring clarity and practical relevance.

Quantitative Data Scientist

Jun 2014 - Aug 2019

Electronica AI – Toronto, ON, Canada

Designed and implemented high-frequency trading algorithms with daily order flow of \$10M+.

- **Trading Algorithm Development:** Collaborated with 5+ engineers to design and implement high-frequency trading algorithms with a daily order flow of \$10M+, leveraging machine learning to improve predictive accuracy.

Quantitative Data Analyst

Aug 2016 - Oct 2018

Alkemi Network – Toronto, ON, Canada

Validated automated financial trading strategies and conducted stress testing on exchanges.

- **Validation and Stress Testing:** Validated automated financial trading strategies based on intra- and inter-exchange price discrepancies and conducted stress testing on exchanges to identify bottlenecks and vulnerabilities in live-like scenarios.

PROJECTS

Wikipedia Article Question-Answering System

Aug 2024 - Sep 2024

NLP pipeline over Wikipedia articles for NER and Question Answering using LangChain, HuggingFace Transformers, and spaCy.

- **Natural Language Processing:** Developed an NLP pipeline over Wikipedia articles for Named Entity Recognition (NER) and Question Answering (QA) using *LangChain*, *HuggingFace Transformers*, and *spaCy*.
- **Similarity Search & Embedding Generation:** Implemented FAISS-based similarity search and embeddings to retrieve relevant text chunks efficiently from large corpora.
- **System Deployment:** Deployed the application using Flask and Docker, with unit tests to ensure reliability.

Natural Language SQL Querying System

Jul 2024 - Aug 2024

Natural language interface that converts user queries into SQL commands using LangChain and HuggingFace Transformers.

- **SQL Generation:** Built an NLP solution that converts user queries into SQL commands using *LangChain* and *HuggingFace Transformers*.
- **API Development:** Implemented RESTful APIs in *Flask*. Used *SQLAlchemy* with *PostgreSQL* for schema management, migrations, and query execution.

Alzheimer's Disease Detection

Jul 2021 - Oct 2021

Deep learning model for Alzheimer's disease detection from MRI scans using TensorFlow and Keras.

- **Deep Learning Image Classification:** Developed a deep learning model for Alzheimer's disease detection from MRI scans using *TensorFlow* and *Keras*. Compared architectures such as *DenseNet121*, *ResNet50*, and *VGG16*.
- **Data Augmentation:** Leveraged datasets including *OASIS* and *ADNI*, with preprocessing and augmentation to improve generalization.
- **Model Optimization:** Applied transfer learning and Bayesian hyperparameter optimization to achieve a notable **10%** accuracy boost in Alzheimer's disease detection.

Coleridge Initiative - Data Citation Extraction [🔗](#)

Mar 2021 - Jun 2021

BERT-based text-to-text transformer model with CRF for extracting data citations. Kaggle competition.

- **Natural Language Processing:** Developed a *BERT*-based text-to-text transformer model in *PyTorch* with CRF for extracting data citations.
- **Data Preprocessing:** Used NLP techniques with *spaCy* and *Gensim* for tokenization, stemming, and stop-word removal.
- **Competition Achievement:** Achieved a top **5%** rank, earning a silver medal in the competition.

Riiid! Answer Correctness Prediction [🔗](#)

Oct 2020 - Jan 2021

Transformer and LightGBM models to predict student answer correctness based on historical performance. Kaggle competition.

- **Time-Series:** Compared *Transformer* and *LightGBM* models to predict whether a student will answer a question correctly based on their historical performance.
- **Model Efficacy:** Attained an AUC score of **80%**, demonstrating strong predictive performance.
- **Competition Achievement:** Achieved a top **3%** position in the Kaggle competition, earning a **silver medal**.

Walmart Sales Forecasting

May 2020 - Aug 2020

LSTM and LightGBM models to forecast 28-day sales for 3,049 Walmart products. Master's thesis project.

- **Time-Series:** Developed *LSTM* and *LightGBM* models to forecast 28-day sales for 3,049 Walmart products using historical and explanatory features.
- **Feature Engineering:** Engineered additional features to capture seasonal patterns and store-level effects, improving forecast quality.
- **Data Analysis:** Analyzed M5 competition data spanning ten Walmart stores across CA, TX, and WI from 2011 to 2016.
- **Optimization:** Applied *Bayesian* optimization to fine-tune model hyperparameters.

Traffic Sign Recognition

Mar 2020 - Apr 2020

CNN to classify German traffic signs achieving 98.14% accuracy on Google Maps Street View images.

- **Image Classification:** Designed and trained a *Convolutional Neural Network (CNN)* to classify German traffic signs, achieving **98.14%** accuracy on Google Maps Street View images.
- **Data Augmentation:** Implemented data augmentation techniques such as zoom, shear, and rotation to improve robustness and generalization.
- **Comparative Analysis:** Compared CNN performance against *ANN*, *Random Forest (RF)*, and *Support Vector Machine (SVM)* baselines.
- **EDA & Model Evaluation:** Performed exploratory data analysis and evaluated models using precision-recall curves, heatmaps, and loss curves.

Genetic Variation of miRNA Primary Transcripts

Apr 2017 - Aug 2017

Investigation of genetic differences in miRNA primary transcripts across human populations. Undergraduate thesis project.

- **Genetic Study:** Investigated genetic differences in miRNA primary transcripts across human populations, analyzing 1,127 SNPs from 1,881 miRNA transcripts to identify key trends.
- **Data Analysis:** Used variance and *Fixation Index (FST)* to quantify genetic differentiation, comparing "average of ratios" and "ratio of averages" formulations for genome-wide FST.
- **Genetic Research:** Showed that African populations exhibit the highest degree of genetic variation, supported by dendrogram-based analyses.

SKILLS

Concepts: Agentic Coding, AI Agent Design & Development, Fullstack & API Development, Evaluation & Testing, CI/CD, RAG, Time-Series Forecasting, Finetuning, Transfer Learning, Computer Vision, NLP

Soft Skills: Technical Leadership & Mentorship, Stakeholder Communication

Languages: Python, SQL, JavaScript, C#, R

Tools & Frameworks: LangChain/LangGraph, Databricks, MLflow, DevOps, Docker, Git, PostgreSQL, Pandas, NumPy, PySpark, PyTorch, TensorFlow, LightGBM, Scikit-learn, HuggingFace, OpenAI API, Azure, Flask, FastAPI

CERTIFICATIONS

BASF 2025 Digital Award – BASF

Jan 2025

Awarded for FOX Talk2Data project - innovation, impact, and enterprise-wide adoption

EDUCATION

Ryerson University – *Master of Science*

Sep 2018 - Aug 2020

Data Science and Analytics – Toronto, ON, Canada

GPA: 4.33/4.33

• **Supervisor:** Professor Konstantinos Georgiou

• **Thesis:** Monthly Sales Forecasting for Walmart with Advanced Machine Learning Techniques

University of Toronto – *Honours Bachelor of Science*

Sep 2013 - Jun 2017

Bioinformatics (Specialist), Computer Science (Major) – Toronto, ON, Canada

Graduated with high distinction; University of Toronto Scholar Award

• Supervisor: Professor Zhaolei Zhang

• Thesis: Detection of Genetic Variations in miRNA Primary Transcripts Using Statistical Analysis