Contact

molakazemi.mj@gmail.com

www.linkedin.com/in/javad-mollakazemi (LinkedIn)

Top Skills

Python (Programming Language)
Machine Learning
Deep Learning

Certifications

Deep Learning for Advanced Computer Vision
Python for Data Visualization
Advanced Python
Databricks Certified Generative AI Engineer
Learning Cloud Computing: Core Concepts

Honors-Awards

Winner of Max Steckler Award (2018)
Winner of travel award of Graduate Student Congress (2020)
1st Place | Global Health Case Competition
Recognized as outstanding PhD candidate in Biomedical Engineering

Javad Mollakazemi, PhD

Senior Data Scientist @ Blueprint Technologies
Boston, Massachusetts, United States

Summary

Analytical and results-driven with over 12 years of experience in leveraging data science, statistics, machine learning, deep learning, and optimization techniques to extract valuable insights and drive informed business decisions.

- Expert at developing and deploying various ML & DL models with full ML lifecycle to solve complex problems and optimize business processes. Excel in multiple programming languages, specializing in data manipulation and preprocessing.

- Certified Generative AI Engineer with expertise in designing and implementing LLM-enabled solutions in using LLM finetuning, RAG, AI agents, vector search, NLP techniques, LLM chains, model deployment and lifecycle management with data governance.

- Team player with strong interpersonal, oral, and written communication skills with bridging the gap between cross-functional teams, and ability to understand complex data and translating to business-driven solutions that drive value and innovation evidenced by completion of over 20 projects, 26 publications, over 12 international conference presentations, and prestigious award as outstanding doctoral candidate.

- Skilled in working with large-scale datasets using distributed computing frameworks ensuring efficient data processing and scalability, and in data visualization to communicate insights to stakeholders.

- Love mountain climbing, hiking, traveling, playing soccer, and outdoor adventures: my cover photo features my climb to mount Damavand, the highest volcanic peak in Asia

Experience

Blueprint Technologies

Senior Data Scientist

July 2020 - Present (5 years 5 months)

• Resolved diverse clients' real-world challenges by executing full ML lifecycle involving end-to-end data pipelines set up, data cleaning, feature engineering, model selection, model evaluation, hyperparameter tuning, and model deployment and monitoring.

• Delivered full-stack GenAI and multi-agent solutions with async coordination, deploying quantized models (4-bit) via Databricks endpoints for cost optimization, integrating web scraping with structured LLM outputs, and building interactive UIs for business stakeholders to access AI capabilities without technical expertise.

• Utilized big data technologies for data preparation, real-time analytics with data streaming, and machine learning such as spark, MLlib, and MlFlow within Databricks and Azure cloud, yielding significant improvements of up to 40% in target metrics.

• Worked on various time-series analytics and forecasting models across different projects and metrics including demand and lead time forecast in supply chain optimization, resource consumption and out-of-resource predictions to optimize cloud usages, resulting in elevated forecasting precision up to 21%.

• Architected production RAG systems with vector databases achieving improved search and retrieval accuracy through optimized chunking strategies, embedding & k-parameter tuning to solve information retrieval challenges.

• Employed Generative AI for tasks such as LLM finetuning and building LLM-powered applications with LangChain, LangFlow, Flowwise, OpenAI SDK, AutoGen, and Hugging Face.

• Optimized cloud workflows including compute configurations, pool, and job optimizations in Databricks resulted in drop in cloud costs up to 80%.

• Designed and crafted advanced dashboards featuring creative data visualization techniques using Power BI and Streamlit App facilitating effective communication of complex data insights to diverse audiences.

University of Kentucky

Data Scientist and Signal Processing Engineer Gained as Graduate Research Assistant

August 2016 - July 2020 (4 years)

Lexington, KY

• Conducted signal processing, data acquisition, and statistical modeling along various data science approaches to investigate the impacts of external stimuli on human bio-signals resulting in publication of 9 research papers.
• Implemented advanced time series analytic methods such as eigenvalues and vectors, Power Spectral Density (PSD), Wavelet time-frequency subcomponents, Principal Component Analysis (PCA), Independent Component Analysis (ICA), and Short-Time Fourier Transform (STFT) to extract insights from signals.
• Served as principal investigator for end-to-end human study, overseeing experimental design and ensuring legal and regulatory compliance by obtaining institutional review board (IRB) permissions and conducting experiments on human subjects.
• Applied various signal preprocessing techniques such as FIR and IIR filtering, discrete Wavelet component removal, signal reconstruction, and Fast Fourier Transform (FFT) to ensure efficient data preparation processes.
• Showcased expertise in product and market knowledge, extensive literature review, staying updated with latest industry trends, and communicating complex topics to diverse audiences evidenced by awards, successful conference presentations, and peer-reviewed publications.

K. N. Toosi University of Technology
Data Scientist Gained as Graduate Research Assistant
April 2013 - August 2016 (3 years 5 months)

• Innovated more than 10 algorithms and designs along conducting detailed analyses on extensive structured and unstructured data, utilizing a variety of data science methodologies resulting in publication of 10 impactful peer-reviewed papers and ranking among top teams in 3 independent coding challenges.
• Applied a range of AI/ML models including ANN, SVM, PNN, MLP-BP, RFC, RNN, RBF with feature evaluation and feature selection techniques such as ANOVA, Chi-Square Test, L1 and L2 regularization, SHAP, FDA and LDA (Fisher’s & Linear Discriminant Analysis) leading to successful completion of 4 projects.
• Performed in-depth data analysis of large bio-signal datasets, utilizing temporal and frequency-based features such as correlations coefficients, PCA, SVD, Fourier and Wavelet transforms, and artificial intelligence models, resulting in publication of 4 research papers.
• Developed novel algorithms in bioinformatics, showcasing expertise in creating new algorithms and efficient design development for fault detections in multimodal data for early detection of various diseases.

• Led multiple projects aimed at predicting and detecting of human diseases through integration of signal processing, pattern recognition, and computer vision techniques with machine learning models, resulting in successful scientific publications and international presentations.

Education

University of Kentucky

Doctor of Philosophy - PhD, Biomedical Engineering · (2016 - April 2021)

K. N. Toosi University of Technology

Master of Science in Control & Dynamical Systems