Mostafa Rezaee, Ph.D.

LinkedIn | Data Scientist | ML Engineer

Ph.D. in Data Science | M.Sc. & B.Sc. in Physics

mostam@bgsu.edu | (419) 315-7481 | Open to relocation.

Summary

- Green Card Holder
- Ph.D. in Data Science
- **ML Engineer**
  Deployed and monitored scalable ML models as REST APIs using FastAPI, Docker, PostgreSQL, Prometheus, and Grafana, with automated CI/CD via GitHub Actions.
- **Data Scientist**
  Developed www.SaveBirds.app, a first-of-its-kind data science web app processing 56 years of data on 800 bird species across 300,000, fueling the $75 billion wildlife-watching industry, reducing analysis time from weeks to minutes, enabling rapid assessment of 40,000+ Protected Areas, and supporting 7 projects with 10 researchers from 6 institutions.
- **AI Scientist**
  Enhanced wearable medical device analytics, improving gait speed accuracy by 25%, step segmentation precision by 57%, step length estimation by 29%, and sway detection accuracy by 46%, while optimizing data processing time by 45% on 1.2M data points.
- **Lead Data Scientist**
  Led the provincial implementation of a nationwide data digitization project for 1,000,000+ students, 40,000 classrooms, and 76,000 teachers across 40 districts, supervising 100+ analysts, improving data accuracy by 40%, enabling real-time updates, and deploying automated alerts for data-driven policy decisions.
- **Publications & projects**
  - Publications (34), including 12 Journal articles & 22 full conference papers.
  - Generative AI in Computer Vision
  - Projects (7) outlined in the SaveBirds.app involving 10 researchers from 6 institutions.
  - All publications and projects are computational studies employing modeling and simulation approaches developed entirely through my own original code, rather than relying on existing software.

Professional Experience

Machine Learning Engineer, Stealth Startup, California, USA *May 2024 – Present*

- Deployed scalable ML models as REST APIs using FastAPI and Docker.
- Implemented real-time monitoring with Prometheus and Grafana.
- Automated ML workflows with CI/CD pipelines.

Data Scientist, SaveBirds, Ohio, USA *Sep. 2019 – May 2024*

- Developed SaveBirds.app to reduce data processing time by 99%.
- Created the Bird Atlas Generator (BAG), making atlas creation accessible without GIS expertise.
- Automated biodiversity metrics for 40,000 Protected Areas.

AI Scientist, Sanofi, Cambridge, MA, USA *Jun. 2022 – Aug. 2022*

- Improved gait speed accuracy by 25% using LSTMs.
- Minimized False Positives in Sway Detection: Improved accuracy in detecting sway by 46% (65% to 95%), reducing false alarms from 22% to 15%. (Technology: Bayesian Filtering, Adaptive Thresholding, Time-Series Anomaly Detection)

- Optimized Data Pipeline: Accelerated preprocessing of 1.2 million accelerometer data points, reducing processing time by 45% (9.2s to 5.1s). (Technology: Apache Spark, Feature Engineering)

Lead Data Scientist & Deputy Director of Research and Technology, Farabi Inst. *Sep. 2013 – May 2022*

- Led the provincial implementation of a nationwide data digitization project, transforming the educational ecosystem for over 1,000,000 students, 40,000 classrooms, and 76,000 teachers across 40 districts.
- Supervised a team of 100+ data analysts directly reporting to me, ensuring standardized, high-quality data collection and analysis at scale.
- Enhanced data accuracy by 40% and established real-time updates, allowing instant visibility into changes in student, teacher, and school profiles.
- Applied advanced regression, classification, and time series analyses to derive actionable insights, guiding data-driven policy decisions for senior authorities
- Developed an automated alert system that identified significant performance shifts, prompting timely interventions and continuous improvement throughout the education system.

Education & Skills

- **Data Science, Ph.D.**, Bowling Green State University, Ohio, USA *Sep. 2019 – May 2024*
- Physics, M.Sc., Focusing on Nonlinear Optics, specifically thermal effects in Green Lasers.
- **Skills:** Computer Vision, NLP, LLMs, RAG, Python (TensorFlow, PyTorch, Scikit-learn, Pandas, NumPy, ipywidgets, GeoPandas), R, SQL, C++, FORTRAN, MATLAB, High-Performance Computing (HPC), Apachi, Hadoop.