

## Ph.D. in Data Science

GitHub: [github.com/0-mostafa-rezaee-0](https://github.com/0-mostafa-rezaee-0)

LinkedIn: [linkedin.com/mostafa-rezaee](https://www.linkedin.com/company/mostafa-rezaee)

Email: [mostafa.mohammadrezaee@gmail.com](mailto:mostafa.mohammadrezaee@gmail.com)

Mobile: 4193157481

- **Placeholder:** Placeholder Placeholder Placeholder Placeholder Placeholder Placeholder Placeholder Placeholder Placeholder Placeholder Placeholder Placeholder Placeholder
- **Placeholder:** Placeholder Placeholder Placeholder Placeholder Placeholder Placeholder Placeholder Placeholder Placeholder Placeholder Placeholder Placeholder Placeholder Placeholder Placeholder Placeholder Placeholder
- **Placeholder:** Placeholder Placeholder Placeholder Placeholder Placeholder Placeholder Placeholder Placeholder Placeholder Placeholder Placeholder Placeholder Placeholder Placeholder Placeholder Placeholder Placeholder
- **Publications:** Placeholder Placeholder Placeholder Placeholder Placeholder Placeholder Placeholder Placeholder Placeholder Placeholder Placeholder Placeholder Placeholder Placeholder Placeholder Placeholder Placeholder
- **GitHub:** Curated and maintained an extensive portfolio of repositories demonstrating technical expertise and open-source contributions across diverse domains, including LLMs, AI Engineering, ML Engineering, Recommender Systems, Causal Inference, A/B Testing, Physics & Engineering, and Personal Branding.

# Machine Learning Engineer

*May 2024 - Present*

- Deployed scalable ML models as REST APIs using FastAPI and Docker, integrating PostgreSQL for persistent prediction storage.
- Implemented real-time monitoring with Prometheus and Grafana, tracking API performance, model predictions, and system health.
- Automated ML workflows with CI/CD pipelines, enabling seamless model updates, containerized deployments, and continuous integration using GitHub Actions.

*SaveBirds*

*Sep 2019 - Apr 2024*

- 99% reduction in data preparation and analysis time—cutting it from 90 days to just 1 minute—by developing the SaveBirds.app, enabling ecologists and conservationists without coding skills to access 56 years of 800 bird species from 300,000 locations across North America.
- 99% reduction in atlas creation time—cutting it from 180 days to under 10 hours—by creating the Bird Atlas Generator (BAG), making it accessible without advanced GIS expertise.
- Automated key biodiversity metrics calculation enabling rapid, data-driven conservation decisions for 40,000 Protected Areas and supporting the \$75 billion wildlife-watching industry.
- Supported 7 projects involving 10 researchers from 6 institutions.

*Sanofi*

Jun 2022 - Aug 2022

- **Boosted Gait Speed Accuracy:** Improved prediction accuracy from 70% to 88% (a 26% relative improvement) using LSTM networks on accelerometer data.
- **Refined Step Segmentation:** Improved classification accuracy from 86% to 94% reducing error rates from 14% to 6% (a 57% relative improvement) through Random Forest, SVM, and advanced signal processing techniques.
- **Reduced Drift in Step Length Calculation:** Enhanced stride length estimation accuracy from 75% to 97% (a 29% relative improvement) and reduced measurement error from 6.5 cm to 5.1 cm using advanced Kalman filtering techniques.
- **Minimized False Positives in Sway Detection:** Improved accuracy in detecting sway from 65% to 95% (a 46% relative improvement) and reduced false alarms from 22% to 15% using Bayesian filtering, adaptive thresholding, and time-series anomaly detection algorithms.

- **Optimized Data Pipeline:** Accelerated preprocessing of 1.2 million accelerometer data points, reducing processing time from 9.2s to 5.1s (a 45% relative improvement) by implementing Apache Spark and advanced feature engineering techniques.

**Lead Data Scientist & Deputy Director of Research and Technology**

*Farabi Institute* Sep 2013 - Aug 2019

- 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.

**Adjunct Professor of Machine Learning and Computational Studies**

*Payame Noor University (PNU)* Sep 2013 - Aug 2019

- Empowered 600+ students over six years by teaching a 3-credit Computer Programming course each semester, covering Python and Machine Learning to 50+ senior and master’s students per class.
- Supervised 100+ senior students’ final projects, designing specialized Machine Learning projects that laid the foundation for their master’s research and careers in AI.
- Shaped the future of AI professionals by inspiring students to integrate Machine Learning into their master’s projects, fostering a new generation of ML practitioners and researchers.

**EDUCATION**

---

**Data Science, Ph.D.** | *Bowling Green State University, Ohio* 2024

**SKILLS**

---

**Data Science** A/B Testing, Causal Inference, Recommender Systems  
**LLMs & Generative AI** Transformers, RAG, LangChain, OpenAI, Claude, Prompt Engineering, Ray  
**AI Engineering** FastAPI, Streamlit, Gradio, HuggingFace, ONNX Runtime, Triton, W&B  
**ML Engineering** Docker, Kubernetes, MLflow, Kubeflow, Apache Airflow, SageMaker  
**Deep Learning** Computer Vision, Natural Language Processing (NLP), HuggingFace, PyTorch, TensorFlow  
**Machine Learning** scikit-learn, caret, XGBoost, LightGBM, Random Forest  
**Programming Languages** Python, SQL, R, C++, FORTRAN  
**Statistics** Hypothesis Testing, Bayesian Methods, Experimental Design, Probability Theory  
**Database** PostgreSQL, MySQL, MongoDB, Redis, Elasticsearch  
**DevOps/MLOps** CI/CD Pipelines, GitHub Actions, Jenkins, Monitoring (Prometheus, Grafana)  
**Cloud Platforms** AWS, Azure, Google Cloud  
**Time Series Analysis** Transformers, TCN, Prophet, LSTM, Statsmodels, ARIMA, SARIMA, TSA  
**Big Data** Apache Spark, Hadoop, Apache Hive, Presto, Apache Flink, Dask  
**Software Engineering** Git, Design Patterns, Test-Driven Development, pytest, Unit Testing  
**Visualization** Matplotlib, Seaborn, Plotly, Tableau, PowerBI