

Reza V. Mehrizi, PhD

Data Scientist, Machine Learning Developer

PROFILE [\(Click here\)](#)

With over 8 years of experience in data science and leading end-to-end machine learning workflows, I'm passionate about translating data into actionable insights that drive product success. My Ph.D. in Statistics, combined with my strong understanding of marketing and quantitative finance, empowers me to build and refine predictive models to inform product decisions, analyzing user behavior to identify opportunities for improvement. I possess a passion for innovation, blending traditional marketing techniques with modern machine learning approaches to develop cutting-edge solutions.

Highlights of Qualifications:

- Proficient in **Python**, including data processing using **Numpy, Pandas, PySpark**, data visualization using **Matplotlib, Seaborn**, ML models deployment using **scikit-learn, TensorFlow, Keras, PyTorch**.
- Adept at **data visualization**, including creating interactive visual dashboard using **Tableau, Power BI**.
- Proficient in working with both **SQL** and **NoSQL** databases using **MySQL, MongoDB, Snowflake**.
- Proficient in handling **Big Data** and **distributed computing** using **Apache Spark, Hadoop, Kafka**.
- Skilled in implementing **MLOps** practices using **Docker, Kubernetes**.
- Proficient in **computer vision** deployment with **HuggingFace transformers** such as **YOLOv8, DETR**.
- Proficient in **NLP** and **LLM** using advanced transformers such as **GPT-4, BERT, Mistral-7B, llama2**.
- Proficient in **AWS**, with expertise in data services including **Amazon S3, Redshift, Glue, Amazon EMR, Data Pipeline**. Skilled in AWS's ML/AI tools like **Amazon SageMaker, Rekognition**.
- Collaborated with a diverse range of industries, including **General Motors, the Ontario Ministry of Health, Expedia, Rogers, and the Ontario Ministry of Transportation**.
- Proven ability to collaborate with cross-functional teams (scientists, engineers, product managers) and translate technical findings into actionable insights for stakeholders.
- Proven track record of successfully leading and delivering complex, scalable data projects.
- Demonstrated leadership by mentoring junior data scientists providing support and technical guidance.

PROFESSIONAL EXPERIENCE

Data Scientist at [MVS Lab](#)

October 2021 – Present

- **Autonomous Shuttle Bus:** [\(Click here\)](#) Collaborated with General Motors to develop cutting-edge ML/AI technologies for an autonomous shuttle bus, incorporating computer vision (OpenCV, TensorFlow), AI predictive models (scikit-learn, TensorFlow), and NLP capabilities (HuggingFace Transformers). These advancements significantly enhanced the vehicle's visual perception and overall operational efficiency.
- **Cable Robot Operation:** [\(Click here\)](#) Developed a cutting-edge CableBot system, integrating ML (scikit-learn, TensorFlow) and graphical (networkx) models for automation in supply chain operations.
- **Warehousing Control System:** [\(Click here\)](#) Designed and implemented an end-to-end warehousing control system using ML (TensorFlow, PyTorch) and AI techniques (intelligent decision-making), resulting in a 17% surge in warehouse throughput efficiency.
- **Health Monitoring System:** Created a holistic health monitoring system for vehicles in collaboration with General Motors using ML/AI predictive maintenance, achieving significant cost saving.
- **Anomaly/Fraud Detection:** Developed a Deep Learning and Bayesian Networks-based Anomaly/Fraud detection algorithm (scikit-learn, PyTorch, PyAgram) for automotive industry, leading to substantial cost savings and enhanced reliability.

Research Assistant at [University of Waterloo](#)

September 2016 – August 2021




- **Health Data Anomaly Detection:** [\(Click here\)](#) Collaborated with Ontario Ministry of Health to develop a highly accurate anomaly detection algorithm using ML predictive models (scikit-learn, TensorFlow) for health data, enabling precise disease diagnosis and prediction in the healthcare system.
- **Pattern Recognition:** Collaborated with Expedia to develop ML predictive (PyTorch, TensorFlow) and generative models (NLTK, Transformers), resulting in remarkable enhancements in productivity, service quality, and cost-effectiveness.
- **Statistical Consultation and Data Analytics:** Guided students, faculty, and industry professionals in overcoming data science and engineering challenges, demonstrating my strong ability to communicate complex concepts in cutting-edge research projects.

Faculty Member at [Semnan University](#)

September 2010 – August 2016

- **Bank Ranking System:** Developed a predictive model for forecasting price fluctuations in the oil industry using statistical time series models and ML techniques (scikit-learn, TensorFlow).
- **Oil Price Forecasting:** Developed a predictive model using time series and ML techniques (scikit-learn, Statsmodels) to forecast price fluctuations in the oil industry.

Detail

-  Toronto, Ontario, Canada
-  (+1) 226-989-5194
-  valiollahi.reza@gmail.com

Links

-  [Website](#)
-  [Google Scholar](#)
-  [LinkedIn](#)
-  [GitHub](#)

Skills

• Statistical Analysis

Bayesian Statistics Anomaly Detection
Graphical Models Time Series Analysis
Experimental Design A/B Testing

• Machine Learning

Predictive Modeling Neural Networks
Deep Learning/Reinforcement Learning (TensorFlow, PyTorch, Keras, PySpark)
Data Pre-processing/Cleaning (scikit-learn)
Visualization Tools (Tableau/Power BI)
Data Mining / Pattern Recognition
Web Scraping (BeautifulSoup, Scrapy)
Ensemble Methods (RandomForest, Boosting, XGBoosting)

• NLP / LLM

Natural Language Processing (Pipelines/ Transformers in HuggingFace)
Large Language Model (Langchain, GPT)

• Programming Languages

Python (Pandas,Numpy, scikit-learn, requests, sqlite3, Django, NLTK)
SQL/NoSQL Databases (MySQL, MS SQL)
API Development (Git, Docker)
Data management (Databricks, Airflow, Snowflake)

• Cloud Computing

Data Warehousing & ETL (AWS, Azure)
AWS (Glue, RedShift, Sagemaker, EC2)
Azure (Databricks, CI/CD pipelines)
Google Cloud (BigQuery)

• Interpersonal Skills

Effective Communication
Mentoring/Advising



PROJECTS

Big Data Projects:

- **Big Data Analytics for Expedia:** Leveraged *Apache Spark*, *Hadoop* and *Hive* to analyze massive Expedia data sets on cloud platforms like *Snowflake* and *Azure*, extracting valuable insights for the travel industry.
- **Vehicle Health Monitoring System:** Designed and implemented a big data pipeline using *Apache Spark* and *Hadoop* to ingest and analyze a vehicle health data lake on *AWS (S3)* and *Azure (Blob Storage)*.

Cloud Service Projects:

- **Azure Cloud Architect for Warehousing Control System:** ([Click here](#)) Designed and implemented a data-driven warehousing control system leveraging data pipelines on *Azure Data Factory*, *Databricks* and predictive models on *Azure Machine Learning Studio*.
- **AWS-Powered Vehicle Health Monitoring System:** Developed a comprehensive health monitoring system for the automotive industry on *AWS*, leveraging machine learning techniques on *Amazon SageMaker* and *Amazon Machine Learning*.

Business Intelligence Projects:

- **Automated Warehousing Control System:** ([Click here](#)) Designed and implemented a data-driven control system for a next-generation warehousing solution, using *ML* and *AI* algorithms to optimize task allocation, path planning, and real-time decision-making for efficient and intelligent warehouse operations, resulting in a 17% increase in throughput efficiency.
- **Computer Vision for Autonomous Vehicles:** ([Click here](#)) Leveraged *ML* and image and video processing techniques like *YOLOv8* and *Faster R-CNN* to develop a robust computer vision system for autonomous vehicles at General Motors. This system contributed to a significant 14% improvement in accuracy, enhancing the overall capabilities of self-driving cars.
- **Data Analysis of AI Trends on YouTube:** ([Click here](#)) Leveraged web scraping techniques with *Google API* credentials to gather YouTube data concerning Artificial Intelligence. The project involved comprehensive data structuring using *SQL*, Python libraries *scikit-learn* and *NLTK*. This process yielded valuable insights into evolving trends within the YouTube AI content landscape.

Fraud Detection Projects:

- **Expedia Fraud Detection:** Implemented a machine learning-based system for fraud detection at Expedia. This system analyzed customer data and transaction patterns to identify and prevent fraudulent activities, resulting in a remarkable improvement in detecting fraudulent activity.
- **Health Data Anomaly Detection:** ([Click here](#)) Developed a *ML/AI* based anomaly detection system using *ML* techniques to identify unusual patterns in health data.
- **Proactive Vehicle Fault Detection:** Collaborated with General Motors to develop a data driven system for proactive fault detection in vehicles. This system facilitated preventive maintenance, leading to a 9% reduction in maintenance costs.

Image and Video Processing Projects:

- **Computer Vision for Autonomous Vehicles:** ([Click here](#)) Developed a robust computer vision system for autonomous vehicles utilizing cutting-edge models *YOLOv8* and *Faster R-CNN*. This system enhanced the vehicle's ability to perceive its surroundings, including pedestrians, vehicles, and obstacles.
- **Interactive Object Detection Platform:** ([Click here](#)) Built a user-friendly web application for real-time object detection and tracking in images and videos. The platform leverages Python's *OpenCV* library and integrates *YOLOv8* for powerful object identification.

NLP and LLM Projects:

- **Interactive Video Analysis Platform:** ([Click here](#)) Developed a web application using the power of *NLP* and *LLMs* from *HuggingFace* and *Langchain*. Users can receive a comprehensive summary of a video along with content analysis, and even ask interactive questions through a chatbot interface powered by the *GPT4*, *Mistral 8x7B* and *Llama2*.
- **Interactive In-Vehicle Passenger Platform:** Designed an in-vehicle passenger communication platform leveraging *NLP* and *LLMs*. This system enables passengers to interact with the vehicle through voice commands.



EDUCATION

PhD in Statistics, [University of Waterloo](#), Ontario

September 2017 – August 2021

Masters in Statistics, [University of Waterloo](#), Ontario

September 2016 – August 2017



SELECTED PUBLICATIONS ([Google Scholar](#))

- Shu, K., **Mehrizi, Reza V.**, Li, S., Pirani, M., & Khajepour, A. (2023). Human Inspired Autonomous Intersection Handling Using Game Theory. *IEEE Transactions on Intelligent Transportation Systems*.
- Sun, C., Cui, Y., Đào, N. D., **Mehrizi, Reza V.**, Pirani, M., & Khajepour, A. (2023). Medium-Fidelity Evaluation and Modeling for Perception Systems of Intelligent and Connected Vehicles. *IEEE Transactions on Intelligent Vehicles*.
- **Mehrizi, Reza V.**, and Shojaeddin Chenouri. "Valid post-detection inference for change points identified using trend filtering." *arXiv preprint arXiv:2104.12022* (2021).