Reza V. Mehrizi, PhD

Data Scientist, Machine Learning Developer

PROFILE (Click here)

With 8+ years of experience as a data scientist and ML developer, I specialize in implementing end-to-end ML workflows. Having a Ph.D. in Statistics, I possess a strong foundation in data science principles. My dedication lies in exploring advanced statistical and ML approaches, leveraging AI techniques, and optimizing ML workflows to architect robust and efficient solutions.

Highlights of Qualifications:

- Proficient in **Python**, including data processing using ***Pandas, PySpark***, data visualization using ***Matplotlib, Plotly*** and ML models deployment using ***scikit-learn, TensorFlow, Keras, PyTorch***.
- Adept at **data visualization**, including creating interactive visual dashboards using ***Tableau, Power BI***.
- Proficient in working with both **SQL** and **NoSQL** databases using ***MySQL, MongoDB, Snowflake***.
- Proficient in handling **Big Data** and **parallel/distributed computing** using ***Apache Spark, Hadoop***.
- Skilled in implementing **MLOps** practices using ***Docker, Kubernetes***.
- Proficient in **computer/machine vision** deployment with *HuggingFace transformers* such as ***vit-base-patch32, ResNet, BEiT, YOLOv7.***
- Proficient in **NLP** and **LLM** using advanced transformers such as ***GPT-3.5, BERT***, and the latest models like ***RoBERTa, Mistral-7B-v0.1, Starling-LM-11B-alpha***.
- Proficient in **Azure**, with expertise in data services including ***SQL Database, Data Lake and Blob Storages, Synapse Analytics***. Adept at automation workflows using ***Azure Logic Apps, Data Factory***. Skilled in Azure's ML/AI tools like ***Azure Machine Learning, Cognitive Services, Databricks***.
- Collaborated with diverse industries including *General Motors*, *Ontario Ministry of Health* and *Ontario Ministry of Transportation*, engaging with stakeholders from executives to Data Science leadership.
- 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, University of Waterloo

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/machine vision, AI predictive models, V2X communication, NLP/LLM, and predictive diagnosis and maintenance systems. These advancements significantly enhanced the vehicle's visual perception and overall operational efficiency.
- **Cable Robot Operation:** (Click here) Developed a cutting-edge cable robot system, integrating ML and graphical models for automation in supply chain operations.
- **Warehouse Control System:** Designed and implemented the end-to-end development of a warehousing control system, orchestrating data engineering processes, Azure cloud infrastructure, and full project lifecycle, resulting in a remarkable 17% surge in warehouse throughput efficiency.
- **Health Monitoring System:** Engineered a holistic health monitoring system for vehicles in collaboration with General Motors using ML/AI predictive maintenance and diagnosis technologies, achieving significant cost savings and heightened vehicle reliability.
- **Anomaly Detection:** Developed a Deep Learning and Dynamic Bayesian Networks-based Anomaly detection algorithm for the automotive industry, resulting in substantial cost savings and enhanced vehicle reliability.

Research Assistant at University of Waterloo

September 2016 – August 2021

- **Covid-19 Anomaly Detection:** (Click here) Collaborated with Ontario Ministry of Healthcare to develop a highly accurate anomaly detection algorithm using ML predictive models for coronavirus, enabling precise disease diagnosis and prediction in the healthcare system.
- **Pattern Recognition:** Collaborated with Expedia Group, an international shipping company, and a sensor fouling system company to develop ML predictive and generative models, 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 GARCH time series and TensorFlow ML techniques.
- **Oil Price Forecasting:** Developed a predictive model using time series and machine learning techniques to forecast price fluctuations in the oil industry.
- **National Census Consultant:** Collaborated on statistical analyses of national census data using A/B/n experimental design for educational and environmental inquiries.

Detail

Waterloo, 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 / Testing
- **Machine Learning**
  - Predictive Modeling
  - Neural Networks
  - Deep Learning/Reinforcement Learning (TensorFlow, PyTorch, Keras, Spark)
  - Data Pre-processing/Cleaning
  - Data Visualization (Tableau/Power BI, Matplotlib/Streamlit)
  - Data Mining / Pattern Recognition
  - Web Scraping (BeautifulSoup, Scrapy)
  - Ensemble/Clustering Methods (Random Forest, Boosting, XGBoosting)
- **NLP / LLM**
  - Natural Language Processing (Pipelines/ Transformers in HuggingFace)
  - Large Language Model (Langchain, GPT)
- **Programming Languages**
  - Python (Pandas, Scikit-learn, Requests, sqlite3, Django/Flask, NLTK)
  - R (ggplot2, tidyr, caret, shiny, forcast)
  - SQL/NoSQL Database Management (MySQL, MS SQL, SQLAlchemy)
  - API Development (Git, Docker)

Cloud Computing

  - AWS (Sagemaker, EC2 instances, Lambda)
  - Azure (Databricks, CI/CD pipelines)
- **Interpersonal Skills**
  - Effective Communication
  - Mentoring/Advising

PROJECTS

Autonomous Shuttle Bus Development for Smart Transportation: (Click here)

- *Object Recognition using machine vision:* Leveraged computer/machine vision like ***YOLO/v7*** and advanced deep learning techniques, to identify and classify diverse objects, including pedestrians, vehicles, and obstacles.
- *Behaviour Prediction:* Developed a robust Behavior Prediction system to foresee the actions of vehicles, pedestrians, and other entities, employing advanced ML and AI approaches.
- *LLM Integration for Enhanced Passenger Interaction*: Implemented ***NLP*** and ***LLM*** techniques to allow passengers to interact with the shuttle bus through voice commands.

End-to-end Warehouse Control System:

- *Azure-Driven Data Architect:* Directed streamlined data workflows, managing data ingestion, transformation, and loading with ***Azure Data Factory***. Enhanced and optimized datasets utilizing ***Azure Databricks*** and ***Synapse Analytics*** for refined efficiency.
- *Azure-Enabled Big Data Analytics and Predictive Modeling:* Utilized ***Azure HDInsight*** for efficient big data analytics and employed ***Azure Machine Learning Studio*** to design and implement predictive models, leveraging graphical modeling services for advanced analytics.
- *Data Visualization and Insightful Dashboards with Power BI:* Utilized ***Power BI*** in Azure to create interactive dashboards and visually represent optimized warehouse performance, making data-driven decisions for enhanced efficiency and strategic planning.

Interactive Video Content Analysis Platform: (Click here)

- *Video Content Analysis:* Leveraged advanced NLP techniques, including the powerful ***Mistral-7B-v0.1*** model, sourced from Hugging Face and Langchain, to extract, transcribe, and summarize video content.
- *Chatbot Integration (LLM-Powered - Llama2):* Integrated an LLM-Powered chatbot using ***Llama2***, derived from ***Langchain***, to enable users to engage in interactive discussions about the video content.
- *Website Development (API Integration):* Designed a website for YouTube video content analysis, creating a user-friendly interface with ***Streamlit*** and incorporating custom APIs.

Real-Time Object Detection and Tracking Using Compute/Machine Vision: (Click here)

- *Object Detection/Tracking Application:* Developed a computer/machine vision application using Python ***OpenCV***, enabling seamless object detection, and tracking in both images and videos.
- *Website Development (API Integration):* Designed a user-friendly website with a Streamlit-based API for real-time object detection and tracking upon upload.

Data Analytics of Artificial Intelligence Trends on YouTube: (Click here)

- *Web Scraping:* Conducted web scraping of YouTube data, employing ***Google API*** credentials, ***Beautiful Soup***, and ***Requests*** libraries for secure and effective data extraction.
- *Data Preprocessing:* Performed data preprocessing with ***SQL***, using ***PostgreSQL*** to comprehensively clean and structure the scraped data.
- *Analysis and Visualization:* Employed Python ***scikit-learn*** and ***NLTK***, to derive valuable insights, along with ***Matplotlib*** **and** ***Plotly*** **I, providing** a deep visual understanding of the evolving trends in AI content on YouTube.

EDUCATION

Doctor of Philosophy in Statistics, University of Waterloo, Ontario

September 2017 – August 2021

- Provided expert statistical consultation and collaborative research support to faculty and industry partners.
- Collaborated on a wide range of projects, offering valuable insights and data analysis solutions that facilitated evidence-based decision-making and problem-solving.

Masters in Statistics, University of Waterloo, Ontario

September 2016 – August 2017

SELECTED PUBLICATIONS

- 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).
- **Mehrizi, Reza V.**, and Shojaeddin Chenouri. "Detection of change points in piecewise polynomial signals using trend filtering." arXiv preprint arXiv:2009.08573 (2020).
- **Mehrizi, Reza V.**, Akbar Asgharzadeh, and Mohammad Z. Raqab. "Prediction of future failures times based on Type-I hybrid censored samples of random sample sizes." Communications in Statistics-Simulation and Computation 48, no. 1 (2019): 109-125.

View More Publications on My Google Scholar.