Location: Montréal area, Québec, Canada
Cell: +1 (438) 838-8856 | **E-mail**: mjt.ghasemi@gmail.com
LinkedIn Profile: https://www.linkedin.com/in/mojtabaghasemi/
GitHub: https://github.com/MJTGhasemi
Personal Site: https://mjtghasemi.github.io/MJ_Portfolio/

Mojtaba Ghasemi

Data Scientist

Summary

A seasoned and results-driven Data Scientist with a Ph.D. in Biomedical Engineering and a graduate degree in Data Science. Skilled in developing advanced analytics and predictive models. Proven ability to leverage advanced analytics, machine learning, deep learning, LLMs, and natural language processing to drive business solutions, strongly emphasizing data querying, cleaning, transformation, and feature engineering. Expertise in explaining complex analytical findings to non-technical audiences and leading cross-functional teams to drive business solutions.

Career Highlights

- Improving data accessibility by 50% by revamping the data visualization dashboard.
- Enhancing stakeholder engagement by 40% by translating complex technical data into clear dashboards.
- Improving AI and data science team proficiency by 45% by fostering a culture of continuous learning.
- Boosting operational efficiency by 30% through cross-functional collaboration & AI Application Development.
- Leveraging advanced AI and NLP techniques, the team demonstrated leadership and problem-solving skills, improving decision-making efficiency by 20%.
- Enhancing machine learning and NLP pipeline efficiency, showcasing adaptability and critical thinking, increased model accuracy by 15%, and reduced processing time by 30%.
- Developing participant safety tracking solutions using R Shiny, integrating them into the data visualization dashboard and company website.
- Securely integrating enhanced dashboards and R components into the company website to provide authorized access for users and sponsors.

List of Qualifications

- ***Programming Languages:*** Python, R, MATLAB, Dax
- ***Data Science:*** Scikit-Learn, TensorFlow, PyTorch, Pandas, NumPy, SciPy, Spark, SQL, NLTK, spaCy, Matplotlib, Seaborn, ggplot2, Sampling, Python Imaging Library (PIL)
- ***Business Intelligence and Reporting:*** Power BI, Tableau, Shiny R, SAP Crystal Report, Meditech, Advanced Excel
- ***Relational Database:*** Microsoft SQL, MySQL, PostgreSQL
- ***Machine Learning:*** Predictive Modeling, Regression, Classification, Clustering, A/B Testing, Deep Learning, Natural Language Processing, Recommendation Systems, Time Series, Association Rules, Model Evaluation, Transformer Models (GPT, Llama)
- ***Advanced AI Techniques:*** Transformer models, Llama, GPT; Fine-tuning LLMs (e.g., GPT-3.5 Turbo), Text generation, Entity recognition, Language model evaluation (BARTScore, CoCo Score, Cosine Similarity), Prompt engineering (zero-shot and few-shot learning)
- ***Machine Learning Pipelines:*** Expertise in designing and implementing end-to-end machine learning pipelines, from data collection and preprocessing to model training, validation, and deployment.
- ***Cloud Platforms:*** Azure, AWS
- ***Containerization Technologies:*** Docker, Kubernetes
- ***Software Engineering:*** Agile, Version Control (Git)
- ***Analytical Skills:*** Data Modeling, Data Visualization, Feature Engineering, Statistical Analysis, Data Analysis, Storytelling
- ***AI Solutions Development:*** Cross-functional collaboration, enhanced data communication, dashboard development
- ***Clinical Research Processes:*** Familiarity with clinical research processes, pharmaceutical industry regulations, and terminology.

- ***Ethics in AI:*** Ethical considerations in AI development and deployment

- ***Soft Skills:*** Communication, Leadership, Problem-Solving, Decision Making, Project Management, Teamwork, Time Management, Adaptability, Business Acumen, Critical Thinking, Team Education

- ***Teaching & Mentoring:*** Curriculum development and teaching data science fundamentals and advanced topics

Professional Experience

ethica CRO – Data Scientist
August 2023 – Present

- **AI and NLP Expertise & ML Pipeline Management:** Demonstrated leadership and problem-solving skills by leveraging advanced AI and NLP techniques, improving decision-making efficiency by 20%. Enhanced machine learning and NLP pipeline efficiency using Python, TensorFlow and PyTorch, showcasing adaptability and critical thinking, which increased model accuracy by 15% and reduced processing time by 30%.
    - Prompt Engineering: Designed and implemented effective prompts for zero-shot and few-shot learning to guide language models in generating Standard Operating Procedures (SOPs) for clinical research.
    - Fine-Tuning LLMs: Fine-tuned GPT-3.5 Turbo, achieving significant improvements in model performance and accuracy for generating SOPs tailored to clinical research needs using transformer models and Pytorch.
    - Evaluation Metrics: Utilized advanced metrics such as BARTScore, CoCo Score, and cosine similarity to rigorously evaluate the fidelity, coherence, and conceptual similarity of generated texts.

- **Cross-Functional Collaboration & AI Application Development:** Exhibited strong communication and teamwork skills by leading cross-functional teams to develop AI solutions, boosting operational efficiency by 30%. Initiated and guided the creation of an NLP product, demonstrating project management and technical leadership.
- **Model Comparison and Selection:** Conducted comprehensive comparisons of different language models (e.g., GPT-3, GPT-3.5, GPT-4, BART, Llama) to identify the best-performing model for generating SOPs.
- **Data Communication & Visualization:** Utilized excellent communication and visualization skills to translate complex technical data into clear dashboards and reports using R and Python (Streamlit, Dash), enhancing stakeholder engagement by 40%. Revamped the company's data visualization dashboard, improving data accessibility by 50%, reflecting strong decision-making and business acumen.
- **Interactive R Statistical Reports:** Developed interactive statistical reports using R and Shiny apps, providing real-time updates and insights for stakeholders.
- **Team Education in AI and Data Science:** Showcased leadership and a commitment to education by mentoring teams in AI and data science best practices, fostering a culture of continuous learning and improving team proficiency by 45%.
- **Dashboard Development for Medical Trials:** Developed complex dashboards using R and Python for sponsors of medical trial experiments to track patient enrollment and progress. Integrated machine learning features and predictive visualizations to provide real-time insights into trial metrics.
- **Project Outcome and Impact:** Successfully generated Standard Operating Procedures (SOPs) for clinical research using GPT-3.5 Turbo, tailored to the specific needs. The project improved SOP generation efficiency and compliance with regulatory standards, contributing to more effective and ethical clinical research practices.

Montreal College of Information Technology – Data Science Program Instructor
October 2023 – Present

- Designed, developed, and delivered the Data Science Program (AEC)
  The Data Science program is designed to equip students with the necessary skills to enter the job market. It includes, but is not limited to:
    - Introduction to Data Science
    - Statistics for Exploratory Data Analysis
    - Data Visualization and Reporting Tool
    - Fundamentals of Python Programming
    - Machine Learning
    - Applied Data Science with Python
    - Advanced topics, including Natural Language Processing and Deep Learning

William Osler Health System – Data Analyst
August 2022 – August 2023

- **Data Analysis & Reporting:** Developed and analyzed comprehensive reports and dashboards of various healthcare data, including HL7, using Microsoft SQL Server, SAP Crystal Report, and Power BI, facilitating efficient decision-making for internal and external stakeholders. Implemented advanced features like tabular models, DAX measures, calculated columns, RLS, and dynamic filters in Power BI.
- **Data-Driven Insights & Approval Achievement:** Translated data insights into strategic modifications using SQL and Python, reflecting critical thinking and business acumen, contributing to earning approval from the Chief Clinical Officer.
- **Data Wrangling in Diagnostic Imaging:** Specializing in data processing and visualization of large datasets from RIS and Meditech, enhancing the understanding of diagnostic imaging data.
- **Machine Learning Model Development:** I constructed machine learning models using Scikit-learn, PyTorch, Pandas, Numpy, Pyodbc, Matplotlib, and Seaborn to predict MRI/CT radiology utilization and optimize inpatient and outpatient services.

Polytechnique de Montreal – Research Associate
January 2016 – January 2022

- **Research Leadership in Cycling Performance:** Guided a team using statistical and machine-learning methodologies to enhance cyclist performance, including developing experimental protocols and ensuring data integrity, fostering a culture of innovation and expertise.
- **Data Analysis and Machine Learning Application:** Utilized advanced data wrangling, clustering techniques (K-Means), and data science methodologies like Recursive Feature Elimination, Principal Component Analysis (PCA), and K-fold Cross Validation for insightful cycling performance predictions.
- **Equipment Data Quality and Database Management:** Established quality metrics for equipment data reliability and accuracy; managed an extensive database encompassing EMG, kinematics, kinetics, and aerodynamics using Python and SPSS.
- **Innovative Technology Development:** Developed an image processing product for biomechanical signal detection using Microsoft Kinect and created a cost-effective, markerless motion capture system, achieving results comparable to high-end systems.
- Mentored graduate students in study design, data analysis, and biomedical innovations.

Musculoskeletal Research Center, Isfahan University of Medical Sciences - Research Associate
April 2013 - March 2015

- Engaging in comprehensive data analysis, using statistical techniques to interpret complex data from force plates and motion capture systems.
- Applying data science and machine learning methodologies to decode patterns, extract insights, and make meaningful predictions from collected data.
- Working closely with diverse patients and athletes, assisting in understanding and improving their health and performance based on data-driven insights.
- Contributing to the development of personalized treatment strategies through advanced data analytics.
- Expanding data science and machine learning expertise in a healthcare environment focused on human biomechanics.

Education

Springboard - Data Science Certificate - 2023

- **Capstone Projects:**
  - Single-cell RNA Sequencing Analysis: Applied machine learning to uncover insights into cellular differentiation and gene-protein interactions (Project GitHub).
  - Insights from Hotel Reviews: Developed a model to predict hotel ratings from customer reviews, enhancing guest satisfaction insights (Project GitHub).
  - Big Mountain Resort Pricing Model: Created a predictive model for ticket pricing to boost revenue (Project GitHub).

Polytechnique de Montreal – Ph.D. in Biomedical Engineering - 2022
Performance Evaluation of Cyclist's Position Using Machine Learning and a Depth Camera-Based Motion Capture System

Amirkabir University of Technology (AUT) – MSc. in Biomedical Engineering - 2015

University of Isfahan – BSc. in Mechanical Engineering – 2013

Certifications

- Azure AI Fundamentals (AI-900) - Microsoft
  - Issued: May 2022 - No Expiration Date
  - Credential ID: 993249642

- Spark NLP for Healthcare Data Scientists - John Snow Labs Training & Certification
  - Issued: Oct 2022 - No Expiration Date
  - Credential ID: 25718612324

- Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans Course on Research Ethics (TCPS 2: CORE)
  - Issued: Sep 2023 - No Expiration Date
  - Credential ID: 0000973495

- Good Clinical Practice
  - NIDA Clinical Trials Network
  - Issued: Sep 2023 – Expires Sep 2026

Publications and Presentations

- Effect of altering different aero handlebar positions on muscle activity, kinematics of the lower limb and oxygen consumption
  Journal of Applied Biomechanics, Submitted, April 2024
  Mojtaba Ghasemi, Daniel Curnier, Maxime Caru, Benjamin Pageaux, Jean-Yves Trépanier, Delphine Périé

- The effect of different aero handlebar positions on aerodynamic and gas exchange variables
  Journal of Biomechanics, 139, p. 111128, May 2022
  Mojtaba Ghasemi, Daniel Curnier, Maxime Caru, Jean-Yves Trépanier, Delphine Périé

- Relation Between Anthropometric Data and Performance of Lower Limb in Squat Jump
  35th Conference of the International Society of Biomechanics in Sports, Cologne, Germany, June 14-18, 2017
  Mojtaba Ghasemi, Ali Mohammadi, Mohammad Ebrahimi, Mohammad Taghi Karimi, Mostafa Rostami

- Nouvelle Approche d'Analyse Aérodynamique de la Position Cycliste Par Méthode CFD
  Congrès 2017 de l'Association Québécoise des Sciences de l'Activité Physique, Université de Sherbrooke, 2017
  Mojtaba Ghasemi, Daniel Curnier, Jean Yves Trépanier, Adrien Koenig, Delphine Périé

- Modeling Electrical Activity of a Neuron: A Bond Graph Approach
  2016 CMBEC39 Conference, Calgary AB, May 24–27, 2016
  Mojtaba Ghasemi, Faezeh Eskandari, Bahareh Hamzehei, Ahmad Reza Arshi

- Crack Detection in Euler-Bernoulli Beams on Elastic Foundation using Genetic Algorithm based on Discrete Element Technique
  Indian J. Sci. Res. 1(2), 248-253, 2014
  Mojtaba Ghasemi, Alireza Ariaei

- Modeling of Bioelectrical Activity of Nerve Cells using Bond Graph Method
  17th International Conference on Biomechanics, Biophysics, and Bioengineering (ICBBB), Vancouver, Canada, 2015
  Mojtaba Ghasemi, Faezeh Eskandari, Bahareh Hamzehei, Ahmad Reza Arshi

Honors and Awards

- Bourse de doctorat en recherche pour étudiants étrangers (FRQNT Fellowship) – 2017-2020
- Bourse de doctorat (Fellowship) – Polytechnique Montréal – 2016-2021
- Bourse compensant les droits de scolarité majorés pour étudiants internationaux – Polytechnique Montréal – 2016-2021
- Master of Science's Full Scholarship - Amirkabir University of Technology – 2013-2015
- Bachelor of Science's Full Scholarship - University of Isfahan – 2008-2013
- Semifinalist in Russian Mathematical Tournaments – 2004