

# Mojtaba Ghasemi

## Data Scientist

[mjt.ghasemi@gmail.com](mailto:mjt.ghasemi@gmail.com) | +1 (438) 838-8856 | [linkedin.com/in/mojtabaghasemi/](https://www.linkedin.com/in/mojtabaghasemi/) | [mjtghasemi.github.io](https://mjtghasemi.github.io)

### Summary

---

Data scientist with a Ph.D. in Biomedical Engineering and a graduate degree in Data Science. Skilled in developing advanced analytics and predictive models, particularly in pricing, promotion, and marketing. Proven ability in data querying, cleaning, transformation, and feature engineering in collaboration with engineering teams. Experienced 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.
- Demonstrating 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, showcasing adaptability and critical thinking, which increased model accuracy by 15% and reduced processing time by 30%.

### Technical Strength

---

- **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:** Tableau | Power BI | Shiny R
- **Reporting:** SAP Crystal Report | Meditech | Advanced Excel
- **Relational Database:** Microsoft SQL | MySQL | PostgreSQL
- **ML Algorithms:** 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 LLM
- **Analytical Software:** IBM SPSS Modeler
- **Software Engineering:** Agile | Version Control (Git)
- **Cloud Platforms:** Azure, AWS
- **Containerization Technologies:** Docker | Kubernetes
- **Analytical Skills:** Data Modeling | Data Visualization | Feature Engineering | Statistical Analysis | Data Analysis | Storytelling | AI and NLP Expertise & ML Pipeline Management | Cross-Functional Collaboration & AI Application Development | Enhanced Data Communication & Visualization | AI Solutions Development
- **Advanced AI Techniques:** Familiarity with state-of-the-art AI models, including Transformer models, Llama, and GPT for building custom large language models
- **Ethics in AI:** Understanding of ethical considerations in AI development and deployment

**Soft Skills:** Self-Driven | Good communication | Leadership | Problem Solving | Decision Making | Project Management | Teamwork | Time Management | Adaptability | Business Acumen | Critical Thinking | Team Education in AI and Data Science

**Teaching and Mentoring:** Curriculum Development for Data Science | Educating in Data Science Fundamentals and Advanced Topics

### Work 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, showcasing adaptability and critical thinking, which increased model accuracy by 15% and reduced processing time by 30%.

- **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.
- **Data Communication & Visualization:** Utilized excellent communication and visualization skills to translate complex technical data into clear dashboards and reports, enhancing stakeholder engagement by 40%. Revamped the company's data visualization dashboard, improving data accessibility by 50%, reflecting strong decision-making and business acumen.
- **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 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.

#### ***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, 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:** Constructed machine learning models using Scikit-learn, PyTorch, Pandas, Numpy, Pyodbc, Matplotlib, and Seaborn to predict MRI/CT radiology utilization, optimizing 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, Ridge and Lasso Regression, clustering techniques, and data science methodologies like Recursive Feature Elimination and Principal Component Analysis 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.
- **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.

#### ***Ariana Movement Galaxy – Data Analyst – March 2015 – June 2016***

- Execution of biomechanical testing, performance modeling, and data analytics.
- Creating custom Python and MATLAB modules for 3D motion capture.
- Data analysis is used to classify vertical jumps (Random Forest, Data Preprocessing).

- Developed Excel dashboards to explain data variation and trends to the non-technical audience.
- Performed complex statistical and prediction reports using Python, R, MATLAB, and SPSS to evaluate professional athletes.

***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**

6-month intensive course in data science, machine learning, visualization tools, Python, and SQL with 1:1 industry mentor oversight.

- **Capstone Project:** Single-cell RNA Sequencing Analysis ([Project GitHub](#))
  - Applied machine learning to uncover insights into cellular differentiation and gene-protein interactions.
- **Capstone Project:** Insights from Hotel Reviews: A Data-driven Approach to Customer Satisfaction ([Project GitHub](#))
  - Developed a machine learning model to predict hotel ratings from customer reviews, enhancing guest satisfaction insights.
- **Capstone Project:** Big Mountain Resort Pricing Model ([Project GitHub](#))
  - Created a predictive model for ticket pricing, significantly boosting potential revenue through data-driven recommendations.

**Polytechnique de Montreal – Ph.D. in Biomedical Engineering - 2022**

Overall GPA: 3.94/4.00

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

**University of Isfahan – BSc. in Mechanical Engineering - 2013**

## Certifications

---

SQL Advanced - HackerRank

Issued: Jul 2023 - No Expiration Date

Credential ID: 657BB19CF368

Python - HackerRank

Issued: Jan 2023 - No Expiration Date

Credential ID: FB8EC0359C1F

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) - Panel on Research Ethics

Issued: Sep 2023 - No Expiration Date

Credential ID: 0000973495

Good Clinical Practice

NIDA Clinical Trials Network

Issued: Sep 2023 – Expires Sep 2026

## 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