MohammadReza **Sadeghian**

Vancouver, BC (Willing to relocate)

(778)682-5124 | mr_sadeghian@sfu.ca | sadeghianmr.github.io | sadeghianmr | sadeghianmr

Summary

Data scientist specializing in transforming complex data into actionable insights. Proficient in using Python, R, and SQL for developing statistical models, machine learning solutions, and interactive dashboards.

Skills

Programming Stacks:	Python (Pandas, NumPy, scikit-learn, Statsmodels, Matplotlib, Seaborn, Plotly) R (tidyverse, ggplot2, dplyr, caret, data.table) • SQL (PostgreSQL, MySQL) • Bash
Statistics & Data Analysis:	Probability & Sampling • Statistical Inference • Hypothesis Testing • Regression Analysis (Linear, Logistic, Poisson/ GLM) • Experimental Design & A/ B Testing
ML & Predictive Modeling:	Classification (K-NN, SVM, Random Forest) • Clustering (K-Means, Hierarchical, DBSCAN) • Neural Networks (ANN, CNN, RNN) • NLP (Topic Modeling, Information Extraction) • Feature Engineering • Hyperparameter Tuning • Model Evaluation
Data Pipelines:	Data Wrangling • ETL/ ELT Workflows • Unstructured/ Text Data Processing • Data Integration (REST API, SQL) • Data Cleaning & Validation (QA)
Visualization & Reporting:	Power BI • Tableau • KPI Dashboard Design • RMarkdown • Quarto
Data Platforms & Tools:	Microsoft Azure • Databricks • RStudio • Jupyter • PySpark • Git • Linux/ Windows

Experience

Data Analyst Intern
IOTO Corporation, North Vancouver

Parliamentary Transcripts Analysis — *May 2025 – June 2025*

- Engineered a **dual-LLM pipeline** to automatically analyze parliamentary transcripts, using a primary model for large-scale **data extraction** and a secondary model for **verification and error correction**, ensuring high accuracy.
- Built an **ideological vectorization framework** combining survey data and transcript analysis to **map MPs' political stances**, cluster parties, and highlight outliers for actionable insights in **public understanding**.

Research Assistant
Simon Fraser University, Burnaby

Hybrid ML-Based System for Plasmid Detection — *Jan 2024 – Apr 2025*

- Developed custom computational tools for plasmid detection, including a **Python/ NetworkX** package for graph-based structural analysis and a hybrid classifier using **Machine Learning** and **Integer Linear Programming**.
- Designed and implemented a **rule-based ensemble system** that integrated outputs from established bioinformatics tools to enhance the overall **accuracy** and confidence in plasmid identification.

Research Intern
Amirkabir University, Tehran

Deep RL for Cryptocurrency Portfolio Management — *Jan 2022 – Aug 2022*

- Collected, cleaned, and engineered large-scale cryptocurrency **time-series data** from **public APIs**; performed **exploratory data analysis (EDA)** to uncover trends, optimize model inputs, and ensure data quality.
- Designed a **Deep Reinforcement** Learning-based portfolio manager in PyTorch to dynamically rebalanced assets, improving Sharpe ratio, returns, and risk metrics compared to baseline strategies.

Education

Master of Science in Applied and Computational Mathematics
Vancouver, BC, Canada

Simon Fraser University — *2024 - 2025*

- Thesis Title: Enhancing Accuracy and Robustness in Plasmid Detection via a Rule-Based Ensemble System

Bachelor's in Computer Engineering
Tehran, Iran

Amirkabir University — *2018 - 2023*

- Thesis Title: A Cryptocurrency Portfolio Management System using Deep Reinforcement Learning Algorithms

Projects

Multi-City Public Transit Performance Analysis
GitHub

INDEPENDENT RESEARCH PROJECT — *Sept & Oct 2025*

- Built an automated Python pipeline to collect and process **live transit data** from multiple public transit agencies and store it in a **PostgreSQL/PostGIS** database designed for efficient spatio-temporal analysis.
- Calculated **key performance indicators (KPIs)** including on-time performance and service frequency using advanced SQL queries, identifying systemic **operational patterns** and performance gaps between different networks.
- Integrated **external datasets**, such as historical weather, to interpret **fluctuations** in transit system performance.
- Developed an interactive **Tableau dashboard** to visualize performance metrics, enabling clear, data-driven comparisons of different urban transit networks.

Age and Gender Prediction from Chest X-rays
GitHub

COURSE PROJECT — *May 2025*

- Developed a multi-task **deep learning pipeline** to predict patient age and gender from chest X-rays using PyTorch.
- Benchmarked various models, ranging from **single-neuron** baselines and **MLPs** to **CNNs** like ResNet and DenseNet.
- Implemented a **Bayesian hyperparameter tuning** with **Optuna** to find the optimal model for maximum accuracy.
- **Interpreted** model predictions by generating **Grad-CAM** heatmaps to identify influential image regions.

Optimized Model Compression for Fine-Tuned Transformers
GitHub

COURSE PROJECT — *July & Aug 2025*

- Engineered an end-to-end **compression and evaluation pipeline** using PyTorch, featuring mixed-precision quantization and intelligent layer filtering to maintain model integrity.
- Benchmarked various innovations against the baseline Delta-DCT method to find the optimal compression strategy.
- Innovated a **data-free compression** method using the discrete wavelet transform, achieving over 5× compression ratio and **30% improvement** over baseline, with negligible accuracy loss on both text and vision transformers.

Toy Company Sales and Trend Analysis
GitHub

PERSONAL PROJECT — *Dec 2024*

- Analyzed **21 months of sales data** to reveal revenue, profit, and seasonal trends across multiple regions.
- Developed an **automated data pipeline** for cleaning, feature engineering, and exploratory analysis, integrating monthly aggregation, profitability metrics, and **category-level KPIs** to evaluate product and regional performance.
- Delivered **business insights** by identifying key products driving revenue and high-profit items, and showing that downtown and commercial stores drove peak holiday sales, supporting inventory and marketing optimization.

Handwritten Digit Recognition
GitHub

COURSE PROJECT — *Mar 2021*

- Implemented a fully connected **artificial neural network (ANN)** from scratch using NumPy to classify handwritten digits on the MNIST dataset.

Volunteer Experience

Head Teaching Assistant
Amirkabir University, Tehran

ACADEMIC SUPPORT AND COURSE COORDINATION — *Feb 2021 - Dec 2023*

- Volunteer as Head TA across multiple courses, helping students understand core concepts through office hours, lab sessions, and clear learning resources to actively support the student community.
- Managed and organized various tasks, including recreating lecture slides, designing lab manuals, exercises, and projects for over 100 students, strengthening my leadership, communication, and organizational skills.

Robotics Mentor
IranKitro, Tehran

COACH FOR YOUTH ROBOTICS COMPETITION TEAMS — *Oct 2018 - Mar 2018*

- Coached and prepared youth teams for national and international robotics competitions, including the Asia-Pacific RoboCup (Junior Rescue Maze League) and FIRA RoboWorldCup Open (Mission Impossible U14).

Certifications

ONLINE COURSE

Data Scientist Associate Certification, 2025 — *DataCamp*

Decision Optimization with Classical and Modern AI, 2025 — *Instats*