

Nima Sarajpoor

Github: <https://github.com/NimaSarajpoor>

LinkedIn: <https://www.linkedin.com/in/nimasarajpoor1991/>

Ontario, Canada

nimasarajpoor@gmail.com

+1 (587) 284 8089

About me

With a PhD degree in Engineering, specializing in data mining, I possess a robust foundation in mathematics and statistics. Currently serving as a full-time Data Scientist at Financial Corporation Manulife, I contribute my expertise to drive insightful data-driven solutions. In addition, I am a key contributor to STUMPY, an open-source data analysis tool, showcasing my commitment to advancing analytical capabilities. My versatile skill set encompasses proficiency in Python, SQL, Shell scripting, and Git, coupled with a profound understanding of Machine Learning, both in supervised and unsupervised contexts. I am experienced in cloud computing, critical thinking, refactoring and unit testing. Beyond technical aptitude, I bring a keen business acumen to align data strategies with organizational goals.

Highlights (hyperlinks included)

• Data Science and System design (Fraud Detection Section)

- **Developed business-driven classifier for risk management**, added unit tests and data quality checks, worked on new Cloud platform, maintaining relationships with the business, conducted several sessions to gain a thorough understanding of the problem before delving into the solution. Coordinated data engineers, data architects, machine learning engineers to distribute the workload efficiently and ensure each team member contributed based on their strengths.
- **Enhanced and maintained 3+2 projects on Fraud Detection**, Took responsibility for maintaining and enhancing three main projects and provided support for two additional projects on an intermittent basis. Addressed low data quality issues by redesigning the analytic process. Proposed a new system design to enhance the fault tolerance of the project and conducted a thorough review of the code base to identify potential opportunities for reducing false positives.
- **Cloud platforms**, Supported both the Machine Learning Engineers and the consulting company in migrating projects from one Cloud platform to another. Provided clear guidelines and coordinated the teams, sharing my knowledge to expedite the troubleshooting process and meet tight deadlines.
- **Worked on new ideas**, On the side, worked on setting up the first GitHub Actions workflow to run tests with each push, ensuring prompt detection of any faulty code. Experimented with fetching metadata from the project's GitHub repository to explore opportunities for establishing a unified process for reporting and tracking issues.
- **Maintained a good relationship with my teammates and the business users**, Known for flexibility, I adapted quickly to new environments and fostered constructive relationships with my colleagues, contributing to an enhanced sense of teamwork.

• Machine Learning

- **Experienced with several libraries in Python**: TensorFlow, Scikit-learn, Numpy, Pandas, Numba, and spark.
- **Experienced with supervised and unsupervised machine learning models** (see my journal and conference papers in my Google Scholar

- **Developed support for feature group in the machine learning tool MLXTEND** to enhance feature selection module, particularly for categorical data (see CHANGELOG)
- **Implemented arbitrary-length anomaly detection algorithm** while reducing its computing time by providing a warm start (see MERLIN)
- **Developed a major feature in STUMPY tool to support top-k matching patterns in data** (see PR#595 and PR#586)
- **Created RandomForestExplorer package** to extract frequent rules from decision paths of RF
- **Created tutorials in machine learning** (see POSTS)

• Programming

- **Proficient in Python**, Invited to join the maintainer community of STUMPY tool, led lectures for a crash course on python with more than 50 participants
- **Strong understanding of data structure and algorithms**, ranked 5th in Programming Contest (CCPC).
- **Experienced with refactoring** (see PR#656), **unit testing** (see PR#657), and **profiling** (see Fast Fourier Transform)
- **Experienced with version control Git, GitHub, and GitHub Copilot**
- **Experienced with Shell scripting** (see Note)
- **Experienced with writing SQL scripts**
- **Familiar with CSS and HTML** (see my blog project)

• Math

- **Strong quantitative reasoning**, verified an equation after fixing the typo (see the long proof here: VALMOD notebook)
- **GPA 4.0/4.0 on university-level courses on mathematics**
- **Took advanced courses:** “data mining [algorithms]”, “applied data science”, “probability and stochastic process”, and “computational statistics.” (GPA: 3.85)

Experiences (in Academia and Industry)

- | | |
|--|---------------------------|
| • Data Scientist in Manulife (Fraud Detection) | Toronto, ON |
| <i>Data Scientist in Group Advanced Analytics</i> | <i>Jan 2023 - Current</i> |
| • Maintainer Collaborator in Stumpy, a data analysis tool | OSS @ GitHub |
| <i>Collaboration at closer level for maintaining the tool</i> | <i>Dec 2023 - Current</i> |
| • [Top] Contributor in Stumpy, a data analysis tool | OSS @ GitHub |
| <i>Developing / Enhancing features</i> | <i>2021 - Dec 2023</i> |
| • Python Lecture Lead at Schulich Ignite | Alberta |
| <i>Python instructor for both beginner and advanced level</i> | <i>2022</i> |
| • Research in data mining area at University of Calgary | UofC, Alberta |
| <i>Developed unsupervised learning techniques for categorizing patterns in energy data</i> | <i>2017 - 2022</i> |

Education

- **University of Calgary** Alberta, Canada
PhD, Dept. of Electrical and Software Engineering (Analyzing energy data) *Grad: Dec 2022*
- **Sharif University of Technology** Tehran, Iran
M.Sc., SUT is ranked 1st among Iranian Universities in Science and Engineering *2013 - 2015*
- **University of Tehran (UT)** Tehran, Iran
B.Sc, UT is ranked second among Iranian Universities in Science and Engineering *2013 - 2015*

Selected Achievements and Awards (hyperlinks included)

- **Full-Stack Machine learning Certificate** Uplimit
Hands-on project for end-to-end process and tracking experiments *2023*
- **Presented a workshop in Toronto Data Workshop** Toronto, ON
Presented a 1-hr workshop on pattern analysis at UofT *2023*
- **Programming Contest (CCPC)** Alberta, Canada
Ranked 5th in Calgary Collegiate Programming Contest *2022*
- **Software Development** OSS
Top contributor in two open-source tools for data analysis STUMPY and MLXTEND *2022*
- **Gordon Lewis Hedberg Doctoral Scholarship** Alberta, Canada
Student with Excellent Grades *2021*
- **Deep Learning Certificates** Deeplearning.ai
Cetificates: DL basics, hyperparameter tuning, CNN, ML project structure in DL *2021*
- **Ph.D. National Competition Award** Tehran, Iran
Ranked 1st among more than 1000 participants of PhD National Exam *2016*
(Afterwards, I applied to Canada.)
- **M.Sc. National Competition Award** Tehran, Iran
Top 10 among more than 5000 participants of M.Sc. National Exam *2013*

Publications

See my publications here: [My Google Scholar](#))

Voluntary Experiences

- **Involved in evaluating candidates in interview** Toronto, Canada
Manulife *2023*
- **Moderator of Employees Engagement Sessions** Toronto, Canada
Manulife *2023*
- **President / Vice President of Grad Students** Alberta, Canada
Department of Electrical and Software Engineering *2017-2020*
- **Evaluator** Alberta, Canada
IEEE award committee, Graduate Student Association *2017-2018*