7025 Belcarra Dr.
Burnaby, BC, V5A 1A6
Email: samira@sfu.ca · Phone: (778) 240-9607

Samir Arora

Curriculum Vitae

Education

Doctor of Philosophy in Statistics *Expected Graduation: April 2029*
Simon Fraser University Burnaby, BC

- CGPA: 4.22/4.33
- Relevant coursework: Statistical Theory I (A+), Linear Models and Application (A+), Generalized Linear Models and Discrete Data Analysis (In progress), Applications of Statistical Computing (In progress), Statistical Consulting I (In progress).
- Research Interests: Computational statistics, Statistical Machine Learning, Monte Carlo methods.

Bachelor of Science in Statistics minor in Computer Science *September 2018 - August 2024*
Simon Fraser University Burnaby, BC

- Statistics CGPA: 4.26/4.33; Overall CGPA: 3.87/4.33
- Relevant coursework: Introduction to Probability and Statistics (A+), Intermediate Probability and Statistics (A+), Introduction to Mathematical Statistics (A+), Programming with C++ (A+), Advanced R for Data Science (A+), Programming with Python (A-), Applied Mutlivariate Analysis (A+), Time Series Analysis (A+), Linear Models in Applied Statistics (A+), Simulation Modeling (A), Introduction to Stochastic Processes (A+), Statistics Theory (A+), Introduction to Artificial Intelligence (A+).
- Relevant projects in courses:
  - *Developed an R package for MARS algorithm*: Developed expertise in the Multivariate Adaptive Regression Splines (MARS) algorithm through an in-depth study of Friedman's 1991 paper. Implemented the MARS algorithm in R and created a comprehensive package accompanied by a detailed vignette for the course "Advanced R for Data Science".
  - *Developed an app providing transit updates*: Utilized the Twitter API to scrape tweets from BC Transit and applied regular expressions to extract relevant information. Developed a user-friendly app that delivers timely updates and information on bus services.

Awards and Scholarships

PhD Research Scholarship (Summer 2025), Graduate Fellowship (Summer 2025), Travel and Research Award (Summer 2025), President's Honour Roll (Fall 2021, Fall 2022), Open Scholarship (Fall 2018, Spring 2019), Gentai Financial Group Undergraduate Scholarship in Finance (Fall 2022), Undergraduate Student Research Award (Summer 2023 and Summer 2024).

Research Experience

Graduate Research Assistant *August 2024 - Present*
Department of Statistics and Actuarial Science, Simon Fraser University Burnaby, BC

- Conducting research on efficient sampling from complex distributions and estimating their normalizing constants under the supervision of Dr. Liangliang Wang.

- Implemented and optimized MCMC algorithms, including parallelization techniques to reduce computational time for large-scale problems.
- Developed proficiency in Julia programming and numerical methods, contributing to reproducible and high-performance research code.

USRA Research Assistant *May 2023 - August 2023 and May 2024 - August 2024*
Department of Statistics and Actuarial Science, Simon Fraser University Burnaby, BC

- Conducted research on the use of pre-trained large language models for solving downstream tasks of summarization and classification under the supervision of Dr. Liangliang Wang.
- Developed a Parameter Efficient Fine-Tuning method for encoder-based BERT Large Language Model that achieves similar (and sometimes even better) performance than full fine-tuning using less than 2% of total parameters.
- Developed skills in Natural Language Processing (NLP), machine learning, neural network architectures (RNN, LSTM, CNN, and transformers), bayesian statistics, and hidden markov models.
- Developed models can be viewed here: https://huggingface.co/Samir001

Work Experience

Teaching Assistant *September 2024 - Present*
Department of Statistics and Actuarial Science, Simon Fraser University Burnaby, BC

- Developed and presented engaging PowerPoint presentations in two tutorials, explaining statistics concepts on a weekly basis to ensure students' understanding.
- Accurately graded weekly quizzes, midterms, and final exams according to the professor's answer key, paying close attention to detail to maintain integrity and detect any instances of cheating or plagiarism.
- Conducted weekly office hours to provide individualized support, giving students the opportunity to ask questions and receive one-on-one assistance.

Strategic Business Analyst / Data Analyst *May 2022 - August 2022*
UniForum Project, Simon Fraser University Burnaby, BC

- Conducted extensive financial data analysis, assessing multi-million-dollar values, to evaluate the efficiency of various university functions for SFU. Conducted comparative analysis with 49 global universities
- Developed a database that handles invoices efficiently and utilized R and Excel to perform data cleaning and transaction classification to predefined codes.
- Applied advanced natural language processing techniques using the `syuzhet` package in R to perform sentiment analysis on SFU employee reviews sourced from platforms like Glassdoor and Indeed.

Publications

1. Syvenky, Z., Miranda, S., **Arora, S.**, Chahma, S. E., Zhang, Z., & Lim, A. (2024). Robot Eye on Your Body Habits: Body-Focused Repetitive Behaviour Detection on Pepper. TSAR Workshop, IEEE RO-MAN 2024.

Talks

1. "Uncertainty Quantification and Model Selection in Neural Networks Using Adaptive Annealed SMC" at International Conference on Statistics and Data Science 2025, Vancouver, Canada.

Personal Projects

Sentiment Analysis of Bitcoin Discussion: Correlation between Lexicon-based Sentiment Scores and Cryptocurrency Price Change

- Performed web-scraping of daily discussion on r/Bitcoin subreddit using `RedditExtractoR` and `tm` packages in R.
- Performed sentiment analysis on all the comments using different lexicons available in `syuzhet` package in R.
- Compared the correlation coefficient between cryptocurrency price change and sentiment scores computed using different lexicons like `afinn`, `bing`, `nrc`, `syuzhet`, and five other lexicons.
- Results: Almost all lexicons had a correlation coefficient of 0.6 except the `bing` lexicon, which had a coefficient of 0.2.

A simulation study to find the best trading strategy over the trading period of ten years

- Simulated a stock exchange with assumptions that the exchange has only one stock and there are only five traders in the exchange, each having there own trading strategy.
- Created a function that simulate stock price data of 10 years using Geometric Brownian Motion. This function takes two parameters to simulate stock prices of stocks of different company sizes and in different economic setting. Created functions for each trader that takes price data as arguments and generate percentage return for 10 years.
- Result: Found that Moving Average Convergence Divergence (MACD) trading strategy performs the best in almost all combinations of economic settings and company size.

Technical Skills and Certifications

- **Programming Languages**: R, Python, Julia, SAS, C++, C, Java, SQL, bash scripting (for submitting jobs on cluster)
- **Software**: MS Excel, MS Access, PowerBI, Tableau, Spark, and Hadoop
- **Certifications**: CFA level 1, LinkedIn Learning Python data analysis, LinkedIn Learning Advanced Pandas, Deeplearning.ai's Deep Learning Specialization
- **Other**: LaTeX, Git