Alexander De Costa

ML Engineer — U of T Math and Stats Graduate
Toronto, Ontario
416 200 4181 alexander.decosta@outlook.com
My current modeling work is confidential; I welcome conversations to share insights and discuss methodologies. Please contact me to learn more.

Objective

ML engineer with a strong background in mathematics and probability, currently building a repertoire of statistical and probabilistic models. Focused on translating cutting-edge research—including Bayesian and kernel methods—into scalable, production-ready APIs. Passionate about bridging theory and practice to deliver real-world AI systems.

Technical Skills

- **Machine Learning:** Built models from scratch (logistic regression, trees, ensembles); developing Gaussian Processes, Bayesian models, and time series techniques
- **Mathematics & Statistics:** Strong foundation in functional analysis, measure-theoretic probability, linear algebra, and inference; currently focused on kernels and probabilistic modeling
- **Tools:** Python, SQL, PyTorch, Jupyter, RStudio, Excel, Power BI
- **ML Infrastructure Tools (in development):** FastAPI, Docker, Kubernetes, MLflow, Airflow, Terraform, Prometheus, AWS

Experience

Manulife *Jan 2023 – May 2023*
Actuarial Student – Experience Analytics

- Updated assumption reports used in valuation and pricing models
- Built data tables and charts using R, SAS, and SQL; collaborated on analytics projects
- Flagged data issues and improved experience study pipelines

Education

University of Toronto *Sep 2020 – May 2025*
BSc, Mathematics and Its Applications (Probability/Statistics)

Projects

Statistical Model Repertoire (Ongoing)

- Building a repertoire of statistical models incl. Gaussian Processes (custom kernels), Bayesian hierarchical models, and time series methods
- Implemented from first principles with a focus on outperforming baselines such as XGBoost and ARIMA
- Engineered for deployment using FastAPI; focused on modularity, scalability, and uncertainty quantification
- Conducting validation and benchmarking on real datasets to ensure reliability