

EDUCATION

Bachelor Of Applied Science, Engineering Science + PEY Co-op | **Machine Intelligence Major** | University of Toronto | Expected June 2027

Relevant Coursework: Data Structures & Algorithms, Mathematics (Probability & Statistics, Linear Algebra), Machine Learning, Databases

Languages: Python, SQL (Postgres, MySQL), MATLAB

Libraries: Pandas, Matplotlib, Numpy, Seaborn, Tensorflow, Keras, Pytorch, Sci-kit, Sklearn, TabPFN, LangChain, Streamlit

Tools: PowerBI, Tableau, Excel, Google Colab, Jupyter Notebook, Docker, Jira, BitBucket, Visual Studio, GitHub, Git

WORK EXPERIENCE

Data Engineer Intern

Aug 2025 - Current

Xylem, Toronto, Canada

- Built and deployed an LLM-powered web app in **Python** with Gemini API, OCR, and Streamlit to automate instrument spec search with **88% accuracy**, adopted by **300+ engineers** enabling them to make faster purchase decisions.
- Integrated Excel MCP server with REST APIs for automated file processing, enabling seamless scaling across workflows.
- Partnered with engineers and software developers to test, debug, and roll out the tool into production, improving cross-team adoption.

Operational Strategy and People Analytics Intern

Jun 2025 - Aug 2025

Amazon - Extern, Remote

- Performed **web scraping** and **sentiment analysis** using **Python** on **135+** Glassdoor employee reviews and **20** YouTube videos to identify operational pain points in Amazon fulfillment centers.
- Created word clouds, graphs, and charts using **Matplotlib** to visualize themes in employee feedback from YouTube and Glassdoor.
- Translated qualitative feedback into structured data and analysis, generating actionable retention strategies.

Data Engineer Intern

May 2024 - Aug 2024

Xylem, Toronto, Canada

- Automated **A/B testing** on **200+** datasets using **Python** and **MATLAB** signal processing toolkit, **reducing manual testing time by ~90%**.
- Used **Excel** and a MATLAB script to automate the calculation of confusion matrix, ROC curves, and determine optimal data threshold.
- Worked in cross-functional teams with a product manager and engineers to optimize the design of tools, **reducing signal noise by ~87%**.
- Wrote detailed statistical reports on testing results, visualizing data through charts to help the engineering team finalize a tool design.

PROJECTS

Advertising Campaign Performance Analytics | [GitHub](#) | Python, Tableau

Aug 2025

- Conducted **A/B hypothesis testing** and **regression analysis** using **Python** on Facebook and AdWords ad campaign analyzing KPIs to demonstrate cost-effectiveness of the superior platform through statistical evidence.
- Created an interactive **Tableau** dashboard to track clicks, conversions, and optimized resource allocation and client ROI, projecting a **30% increase** in conversions and **15% reduction** in costs.

Vendor Performance Analytics | [GitHub](#) | Python, SQL, PowerBI

Jul 2025

- Conducted product analytics to identify underperforming SKUs with high margins for targeted promotional strategy in **SQL** and **Python**.
- Performed **hypothesis testing** on vendor data to confirm profitability differences and validate a **74% unit cost reduction** for bulk purchases.
- Developed an interactive **Power BI** dashboard to track vendor KPIs, identifying **\$2.71 million** in unsold inventory and enabling targeted supply chain interventions.

Data Visualization of Factors Affecting Community Wellness and Solutions for GenWell | [Project](#) | Tableau

Jun 2024

- Created **Tableau dashboards** and used **Excel** to visualize social capital data of **4000+ Toronto residents** from GenWell's database and identified low-income citizens of old Toronto areas to be at greatest risk of social isolation post-pandemic.
- Built a slide deck and presented recommendations to GenWell CEO within 6 hours for the Toronto Community Wellness challenge.

Machine Learning Based Study Habit Recommendation System | [GitHub](#) | [Research Paper](#) | Python

Jan - Apr 2025

- Performed **exploratory data analysis** on **2,392 students'** data using **Python libraries: Pandas, Seaborn, Matplotlib** to identify key features influencing academic performance, guiding effective model training.
- Collaborated in a team of three to develop a personalized academic recommendation system, using TabPFN for grade prediction (achieving an **R² of 0.953**) and kNN with LLM integration using OpenAI GPT-4 API to create text-based study habit suggestions to improve student grades.

LEADERSHIP EXPERIENCE

Vice President of Finance

Jul 2023 - Current

Women in Science and Engineering (WISE), Toronto, Canada

- Conducted data analysis, market research, competitor analysis and co-led a **7-person team** to identify sponsor needs, adjust pricing, and pitch WISE, securing **\$75,500+** in sponsorships in 5 months and a **200% increase in ticket sales** for events with **400+ attendees**.