WILLIAM GLAZER-CAVANAGH

438-389-8650 ◇ willcavanagh@hotmail.com ◇ French / English Website LinkedIn GitHub Scholar

EDUCATION

University of Montréal - Professional Masters in Machine Learning September 2022 - December 2023
Accepted in the MILA program - 3.9 / 4.3 GPA *Montreal, Quebec*

Polytechnique Montréal - Bachelor in Software Engineering September 2017 - May 2022
Artificial Intelligence & Data Science - 3.4 / 4.0 GPA *Montreal, Quebec*

WORK EXPERIENCE

AbCellera Biologics - *Machine Learning Scientist Intern* May 2023 - November 2023

Automation of high-throughput screening (HTS) for drug discovery with online learning and computer vision

- Outperformed pre-existing models in 80% of the images in out of distribution (OOD) settings as measured by A/B testing.
- Enabled processing of several classes of new assays by building a semi-supervised pipeline for automatic labelling with **Python**, **TorchLightning** and **PyTorch**.
- Orchestrated workflows with **Flyte** and implemented distributed data ingestion processes with **Dask**
- Understood and acted on cell screener's issues with existing workflows for a smooth change management

Intact Insurance - *Data Scientist Intern* May 2022 - August 2022

Optimization and evaluation of call center processes across Canada

- Increased yearly call center sales by an estimated 200K$ by implementing employee management strategies in production. These strategies were designed and reviewed using our custom discrete event simulator implemented with **Python**, **SimPy**
- Detected production issues by developing dashboards, ETL pipelines and A/B tests to monitor models with **SQL**, **PowerBI**
- Found and fixed anomalies in our tree-based model by analyzing features and their impact on predictions with **LightGBM**

Croesus Finansoft - *Research Intern* May 2021 - August 2021

Automatically explain stock price fluctuations from news articles with Natural Language Processing (NLP)

- Improved generated text's ROUGE score by 5% by integrating two recently published algorithms to our pipeline using **Python**, **Docker** and **Large Language Models**
- Identified avenues of research by assessing the state of the art in knowledge injection with knowledge graphs for NLP algorithms. Discussed interesting findings with principal researcher to establish a research agenda.
- Increased reproducibility and reduced delays by 15% by refactoring and re-designing the algorithm evaluation pipeline with **Argo Workflows** and **Microservice Design Patterns**

Merck Canada - *Data Engineer Intern* January 2021 - May 2021

Development of Merck Canada's internal ETL (extract, transform, load) processes for Business Intelligence

- Produced first insights on GARDASIL-9 HPV vaccine's by leading a market opportunity analysis. This project is currently being used to power a canada-wide marketing and awareness campaign against HPV.
- Automated Merck Canada's manual ETL workloads and reduced delays by two days with **AWS Cloud** and **Python**

PROJECTS & PUBLICATIONS

IVADO - Deep Learning Essentials Proofread IVADO/MILA online Deep Learning class covering: CNNs, GANs, RNNs

Polytechnique Montreal - MCIS Lab Paper accepted at EMSE: *Towards a Change Taxonomy for ML Systems*

Polytechnique AI Student Club - Founder Kickstarted the club and handled content for the first cohort of 10

ProjectX - Team Lead Research competition to analyze links between respiratory diseases and air quality

Polytechnique Montreal Admission Scholarship Awarded for outstanding academic performance

PROGRAMMING SKILLS

Languages: Python, Bash, C/C++, Java,
Others: Flask, Linux, Docker, GIT, AWS, LaTeX
Books: Clean Code, SWE at Google *(ongoing)*

ETL: Pandas, Dask, SQL/NoSQL
Machine Learning: PyTorch, Lightning, SK-Learn, NumPy
Visualization: MatPlotLib, Seaborne, PowerBI