

Matteo Esposito

SCIENTIFIQUE DE DONNEES ET ETUDIANT A LA MAITRISE AU MILA

514-922-6789 | matteoesposito97@gmail.com | [matteo-esposito.github.io](https://github.com/matteo-esposito) | [matteo-esposito](https://www.linkedin.com/in/matteo-esposito) | [matteo-esposito](https://www.youtube.com/channel/UC...)

Competences

- Programmation** Python, Java, Bash, SAS, R, Git, SQL, \LaTeX
- Librairies** Pandas, Numpy, Scikit-Learn, Pytorch, Torchaudio, SpeechBrain, PyYAML, Shap, HuggingFace, Transformers, Nltk, OpenCV
- Technologies** Google Cloud Compute Engine, Apache Airflow, Google Colab, Github, Gitlab, Linux, Jira, Agile
- Langues** Francais, Anglais, Italien

Experiences Professionnelles Pertinentes

Intact Corporation Financiere

Montreal, QC

SCIENTIFIQUE DE DONNEES I

mai 2020 - Aujourd'hui

- Piloted and presented various proof-of-concepts for expert system implementations in **Python** with Experta and the **Java** Drools framework.
- Maintained a Java **expert system**, managed dependencies and builds with **Maven**, wrote unit tests, monitored **CI/CD pipelines** in Gitlab.
- Created **data preparation pipelines** leveraging **pandas, numpy, SAS and SQL** with **Oracle databases** and automated them using **Apache Airflow**.
- Trained supervised learning models for numerous added functionalities on web quoting service (tailored-web experience).
- Led recurring Agile meetings with the team and stakeholders for project management.
- Coached and supervised 2 data science interns over 8 months.

Aviva Canada

Montreal, QC

ANALYSTE DE DONNEES CO-OP

mai 2019 - avr. 2020

- Evaluated 3rd party geolocation/telematic data source for added value of new features in pricing regression models.
- Developed and maintained an **internal python library** used for actuarial calculations.
- Integrated library with modeling pipeline for premium validation and replaced legacy process. (4x improvement in mean run-time)

Aviva Canada

Montreal, QC

ANALYSTE DE DONNEES CO-OP

jan. 2018 - aout 2018

- Built predictive models using **scikit-learn** to model frequency and severity of personal auto insurance claims.
- Reduced error of **variable imputation** methods in Ontario pricing databases by 20% using **multi-class classifiers**.
- Created a database of at-risk-clients in SAS & trained a **binary classifier** to predict risk of default with 85% out-of-sample accuracy for the optimization of the current risk/client selection process.

Projects

GFlowNets Implementation (Generative Flow Networks)

PART OF PROF. SIMON LACOSTE'S IFT6269 PROBABILISTIC GRAPHICAL MODELS COURSE

- Implemented Generative Flow Network model architecture (GFlowNet) for generative tasks including image generation and hypergrid.
- Work presented at MILA to students and faculty.

Open Source Contribution to SpeechBrain (Concurrent Speaker Counting)

SUMMER PROJECT IN COLLABORATION WITH PROF. MIRCO RAVANELLI

- Create modeling pipeline (data cleaning, augmentation, processing and modeling using **PyTorch**) for XVector and ECAPA TDNN model architectures to identify number of concurrent speakers in an audio sample.
- Synthesize overlapping speech dataset (0-5 speakers) from the **LibriSpeech** audio dataset.
- <https://speechbrain.github.io/>

Education

University of Montreal/MILA

Montreal, QC

MASTERS OF SCIENCE - COMPUTER SCIENCE, MACHINE LEARNING SPECIALIZATION

Sept. 2020 - Jul. 2023 (Expected)

- Professional Masters completed while working full time as a data scientist at Intact Insurance.

Concordia University (Graduation with Distinction)

Montreal, QC

BACHELOR OF SCIENCE - ACTUARIAL MATHEMATICS CO-OP WITH MINOR IN COMPUTER SCIENCE

Sept. 2016 - May 2020

- Completed **4 full time, 2 part time co-op internships & 3 actuarial exams** during studies.
- Member of the Co-op Institute, **Golden Key International Honour Society** (Top 15% of students) & Data Innovation Society.