Matteo Esposito

matteoesposito97@gmail.com | www.matteoesposito.com | matteo-esposito | matteo-esposito

Skills

Programming Python, Java, Bash, SAS, R, Git, SQL, LaTeX
Libraries Pandas, Numpy, Scikit-Learn, Pytorch, Torchaudio, SpeechBrain, PyYAML, Shap, HuggingFace, Transformers, Nltk, OpenCV
Technologies Google Cloud Compute Engine, Apache Airflow, Canada Compute, Google Colab, Github, Gitlab, Linux, Jira, Agile
Languages English (Fluent), French (Fluent), Italian (Intermediate)

Relevant Work Experience

Intact Financial Corporation
Montreal, QC

Data Scientist — *May 2020 - Present*

- 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 an **ETL pipeline** leveraging **pandas, numpy, SAS and SQL** with **Oracle databases** and automated their execution using **Apache Airflow** and make commands. Managed builds and CICD using **Gitlab** and **Docker**.
- Created an internal python package to be used within the ETL pipeline.
- Trained supervised learning models for PoCs for numerous added functionalities on web quoting service (tailored-web experience).
- Coached and supervised 2 data science interns over 8 months.

Aviva Canada
Montreal, QC

Data Analyst Co-op — *May 2019 - Apr. 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

Data Analyst Co-op — *Jan. 2018 - Aug. 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

Total Loss Detection through Images

Final Project in fulfillment of the requirements for Masters in Computer Science

- Implemented **ensemble model and data pipeline** for classification of claim images for purposes of detecting total loss claims using supervised models on tabular data, pretrained **image classification** models on image data and **image segmentation** models from **huggingface** to add information to input data.

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. `https://speechbrain.github.io/`
- Used **Compute Canada clusters** and **GCP Compute Engine** to run, test and debug implementation.
- **Synthesize/create overlapping speech dataset** (0-5 speakers) from the **LibriSpeech** audio dataset.

Education

University of Montreal/MILA
Montreal, QC

Masters of Science - Computer Science, Machine Learning Specialization — *Sept. 2020 - Jul. 2024 (Expected)*

- Professional Masters completed while working full time as a data scientist at Intact Insurance.
- Recipient of the **Mitacs applied research grant** of 30,000$ to fund masters project.

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.