

# Khoi Nguyen

nguyentienkhoi01@gmail.com · (438) 938-0432 · Montreal, Canada

linkedin.com/in/khoi-tien-nguyen · github.com/khoitiennguyen · khoitiennguyen.github.io

## Work Experience

### Intact Insurance, Data Lab

Montreal, Canada

DATA SCIENTIST INTERN

May 2023 – Aug 2023

- Analyzed internal and external data, including client flood claims, historical floods and satellite imagery, using **Python (Pandas, Matplotlib, NumPy)** and geospatial libraries in **R**.
- Utilized and improved weather claims classification model to effectively identify flood claims, revealing key patterns in claim distributions.
- Developed and investigated hypotheses on model efficacy in high-risk areas, contributing to improved understanding of flood risk patterns.
- Consolidated disparate datasets to perform extensive data cleaning and merging using **SQL** and **R**, ensuring data consistency and reliability for model benchmarking.

### McGill University, Distributed Digital Music Archives & Libraries Lab

Montreal, Canada

SOFTWARE DEVELOPER

May 2022 – Apr 2023

- Developed a workflow for user-based image layer labeling and optical recognition model training.
- Implemented image processing functionalities using **Python (NumPy, OpenCV, Django, Flask)**, increasing training data preparation efficiency.
- Collaborated with researchers for iterative modeling changes, leading to significant experimental advancements.
- Designed a dynamic data loading scheduler, tripling model training speed and efficiency.

### Accreon

Fredericton, Canada

SOFTWARE DEVELOPER CO-OP

Sep 2018 – Jan 2019

- Refined **Java**-based **ETL** processes for improved data handling, introducing new features for enhanced flexibility and efficiency in data processing.
- Streamlined client data extraction into bilingual formats and implemented compression techniques, significantly improving data storage and accessibility for developer use in a web application.

## Education

### McGill University

Montreal, Canada

B.Sc. COMPUTER SCIENCE, GPA: 3.53/4.0

Sep 2019 – May 2023

- Awarded the Major Entrance Scholarship for outstanding academic achievement.
- Courses: Data Science, Applied Machine Learning, Deep Learning, Operating Systems, Natural Language Processing.
- Extracurriculars: McHacks Hackathon (2020), McWiCS Hackathon (2022), McGill Orientation Leader (2020 – 2022).

## Skills

**Programming** Python · SQL · R · Java · C · C++ · JavaScript · HTML · CSS · Bash · OCaml · Assembly

**Libraries** NumPy · Pandas · TensorFlow · PyTorch · Scikit-learn · SciPy · Flask · HuggingFace · XGBoost

**Tools** Git · Jupyter · Docker · AWS · PostgreSQL · Spark · Snowflake · Tableau · PowerBI · Airflow · Kafka

## Projects

### Automatic Music Transcription

- Developed an automatic music transcription system, achieving above 90% accuracy in sheet music format conversion using a **CRNN** model adapted for spectrogram image processing.
- Leveraged **Python** for audio processing and data handling, converting MIDI files to spectrogram images and tokenizing musical symbols for efficient model training and evaluation.

### Movie Recommendation System

- Engineered a movie recommendation system incorporating a **Kafka** data pipeline, a **Redis** database, and advanced model training with **XGBoost** and SVD for enhanced accuracy (0.15 RMSE).
- Integrated robust **CI/CD** practices using **GitHub Actions**, ensured load-balancing and system scalability through **Nginx** and **Docker**, and implemented comprehensive monitoring using **Prometheus** and **Grafana**.