

Olivier Nguyen

NLP, Deep Learning, ML Engineering

olinguyen.com

nguyenolive@gmail.com

github.com/olinguyen

EXPERIENCE

ServiceNow Applied Research Scientist

January 2021 — May 2022

Montreal, Canada

- Trained semi-supervised multilingual sentence embeddings for intent classification; improved on all customer benchmarks over previous language models pre-trained on enterprise data
- Led the development of the NLU training pipeline for conversational agents and search that powers hundreds of companies; responsible for migrating from the legacy Java stack
- Trained, evaluated and managed data collection for question answering models deployed for a snippet search feature for the AI Search product
- Developed an NLP framework to quickly pre-train, fine-tune, evaluate and export language models on the internal GPU cluster; library is used by multiple teams to accelerate time from experiments to production

Element AI Applied Research Scientist

September 2018 — January 2021 (acquired)

Montreal, Canada

- Led work on semantic search and question answering for the enterprise knowledge management product; turned proof-of-concepts into core functionalities of the product
- Built and deployed ML models for information extraction from 2D documents for Document Intelligence, a flagship product for automated document processing; led early research efforts, and actively involved in building proof-of-concepts, MVP, and product launch
- Developed and deployed text classification models to accelerate insurance clerks' workflows, reducing the time and number of errors during data entry

NTT Media Intelligence Lab Research Intern

January 2015 – August 2015

Yokosuka, Japan

- Built, trained and evaluated multi-modal models for a media event detection pipeline; published evaluation runs to TRECVID Multimedia Event Detection (MED) 2015
- Ran experiments combining objection detection, deep CNN features with traditional visual and audio features for the MED task submission

PROJECTS

BigScience 🤖

Developing tools for managing datasets for governance and training large multilingual language models [\[Github repo\]](#)

Jeux de Genie '15

1st Place against all engineering schools in Quebec; Developed the software for a fully autonomous robot that had to complete multiple tasks around a complex track [\[video\]](#)

ConSat-3

1st Place against competing Canadian universities; Developed and integrated embedded Linux applications for a CubeSat [\[article\]](#)

SKILLS

Programming languages

- Python
- C/C++
- Java
- Javascript
- Bash
- SQL

ML frameworks & libraries

- PyTorch
- HuggingFace stack
- pandas
- scikit-learn
- NumPy
- spaCy
- TensorFlow/Keras
- MLFlow
- faiss

Software development tools

- Docker
- git
- Flask
- FastAPI
- PostgreSQL

EDUCATION

MASc Elec. & Comp. Engineering

2016 – 2018

University of Waterloo

Advisors: Mark Crowley, Joon Lee

BEng Computer Engineering

2012 – 2016

Concordia University