

Derek Xu

ML Developer

Contact

☎ (647) 636 5430

✉ derek_yhx@hotmail.com

Oakville, ON

Portfolio

<https://dxiled.github.io/>

Education

Bachelor's Degree

Computer Science

Queen's University

Sept. 2018 – Apr. 2023

Kingston, Ontario

Skills

Python (PyTorch,
TensorFlow, NumPy,
pandas, scikit-learn)

Java, C, R

SQL

Version Control (Git)

JIRA and Confluence

Agile Development (Scrum)

AWS

MS Office (Word, Excel)

Strong communication skills

Team player

Languages

Fluent in English

Fluent in French

Work Experience

Data Science Intern at Bell Canada

September 2021 – September 2022 / Hybrid in Toronto, Ontario

- Worked in a small team on natural language processing for a chatbot in both English and French.
- Analyzed and classified production data to evaluate the performance of a chatbot and generate reports.
- Created intents and verified their performance using a Rasa model with QBox testing, reaching 92% training accuracy in English and 87% training accuracy in French.

Projects

Enron Email Spam Classifier

- Worked in a group in an academic setting to develop various ML models for spam classification using the Enron dataset.
- Implemented a custom sub-word-based tokenizer using the byte-pair compression algorithm.
- Achieved a testing accuracy of 85% using a recurrent neural network and the custom tokenizer, performing 10% better than the baseline recurrent model using a word-based tokenizer.
- Achieved a testing accuracy of 97% using a SVM with the custom tokenizer.

Twitter Vaccination Dataset Analysis

- Worked in a group in an academic setting to implement various analytics models to explore and analyse a Twitter dataset about COVID-19 vaccinations.
- Explored a state-of-the-art transformer model for sentiment analysis, achieving similar results to other papers regarding the dataset.
- Implemented K-means clustering, achieving a silhouette score of 0.504 and a Davies-Bouldin index of 0.6.
- Analysed the clusters, revealing a group of tweets that were likely bots made to promote paid vaccinations.

Cancer Nucleus Image Classifier

- Worked in a small group in an academic setting to develop a TensorFlow-based convolutional neural network to detect cancer nuclei in images of tissue samples.
- Worked in a small group in an academic setting.
- Gained a thorough understanding of current methods of image processing and classification, reaching 99.6% training accuracy and 95.8% testing accuracy.
- Explored benefits of tuning parameters, aiming for a higher recall for detecting cancer nuclei.