YILIN (LARRY) LI

(416)-834-8954 ◇ y2432li@uwaterloo.ca ◇ larryli1999.github.io

EDUCATION

University of Waterloo Sep 2017 - Present
B.A.Sc. in Mechatronics Engineering with Artificial Intelligence Option *Waterloo, ON*
Cumulative GPA: 3.92/4.0

PUBLICATION

· *Automatic Classification of Pathology Reports using TF-IDF Features*
Shivam Kalra, **Larry Li**, Hamid R. Tizhoosh
(Arxiv preprint 1903.07406)

RESEARCH EXPERIENCES

University of Waterloo May - Aug 2021
Undergraduate Research Assistant *Waterloo, ON*
· Supervised by Prof. Jimmy Lin, researching on information retrieval and natural language processing
· Developed an information retrieval system with BM25 and T5 model for TREC Misinformation Track
· Experimented Transformer-based models such as BERT, RoBERTa, and DistilBERT on MS MARCO document and passage ranking tasks

University of Waterloo (Kimia Lab) Sep - Dec 2018
Undergraduate Research Assistant *Waterloo, ON*
· Supervised by Prof. Hamid Tizhoosh, researching on medical image search and keyword extraction
· Applied keyword extraction with TF-IDF and LDA from Scikit-learn from 1,949 pathology reports
· Generated over 5,000 patches from 300 pathology images using Openslide Library for DenseNet feature extraction

WORK EXPERIENCES

Huawei Canada Sep - Dec 2020
Machine Learning Engineer Intern *Montreal, QC*
· Applied 8-bit QAT on BERT and fine-tune the fully quantized model on the GLUE benchmark
· Implemented knowledge distillation to stabilize quantization while replacing BERT LayerNorm with NoNorm
· Experimented structured pruning on BERT FFN during model's pre-training and fine-tuning phases

Synapse Technology Jan - Apr 2020
Deep Learning Engineering Intern *Palo Alto, CA*
· Developed a 3D detector with SSD and used the slice-and-fuse architecture to detect handguns from CT scans
· Trained and evaluated a system of SSDs to detect explosives and assembled IEDs from X-ray images
· Implemented automated consensus process on 1.5M labelled data to improve data processing efficiency

Primate Labs May - Aug 2019
Machine Learning Developer Intern *Toronto, ON*
· Developed an Android application which implemented image classification, object detection, semantic segmentation, face recognition, style transfer, pose estimation and sentiment analysis using TensorFlow Lite
· Applied quantization aware training to LeNet which resulted in 4 times size reduction and 30% latency reduction
· Trained MobileNet-V1 on the ImageNet Dataset and achieved 69.1% validation accuracy for image classification

EXTRA-CURRICULUM

University of Waterloo Self-driving Car Team (WATonomous)
Perception Team Core Member
· Improved inferencing performance of TensorFlow neural networks by CPU and FPGA optimizations
· Developed a C++ algorithm for ROS nodes which ran the networks and transferred predictions from the networks to the rest of the software pipeline