PENGNAN FAN

NLP Developer at KingMed Diagnostics

@ pengnan.fan@mail.mcgill.ca +86 130 7102 7025 linkedin.com/in/pengnanfan github.com/Catosine

EXPERIENCE

NLP Developer - KingMed Diagnostics

Medical NLP

Jun 2021 - Now Guangzhou, Guangdong

- Founding member of KingMed NLP Team
- Focusing on BERT-based sentence encoding model development, deployment and inference acceleration
- Core developer of KM-IRS, an BERT-based medical examine result explaining project (backend architecture design/Model Training/Acceleration)

Research Intern - McGill University

Temporal Action Localization on Hockey Videos

Dec 2019 - May 2021 Montral, Canada

- Supervised by Professor Martin Levine from Intelligent Multimodel Video Anaysis Lab at McGill Centre for Intelligent Machines.
- Assisting PhD students on building a hockey video temporal action location dataset (Tempucky)
- Traing BSN/Dual Encoding Network on the Tempucky dataset
- Implementing a video labelling tools in Python3

Machine Learning Intern - Huawei Technologies

Face Identification with Neural Architecture Search

May 2019 - Aug 2019 Beijing, China

- Internship program in summer 2019 at Huawei Beijing Research and Development Centre
- Applying PDARTS/PC-DARTS/Single-Path NAS to search for high efficiency face identification models
- Holding company-wide seminars about AutoML-NAS papers
- Assisting Huawei Noah's Ark Lab on NAS experiments

TECHNICAL SKILLS

- Programming Language: Python, Java, C, CUDA C
- Editor: VS Code, Eclipse, PyCharm
- ML: PyTorch, Numpy, Pandas, Faiss, OpenCV
- Backend: Flask(Python), JavaSpring(Java)
- Cloud: Docker, Kubernetes
- Software Engineering: UML, Umple, Git, Gherkins, JUnit
- Graphics: OpenGL(Java-based)

EDUCATION

Bachelor of Software Engineering

Department of ECE, McGill University

January 2018 - May 2021

CGPA: 3.70/4.0

PROJECTS

KMIRS: Medical Examine Report Explaining System

- A BERT-based Chinese medical examine explaining system
- Implemented a encoding micro-service to encode Chinese/English input with different language model
- Designed and implemented a matching subsystem to match incoming microorganism results with pre-defined explanation from database
- Implemented a test script for the system

Video-Text Retrieval System

- 4-member design project supervised by Professor Martin Levine at McGill
- Developed the system based on *Dual-encoding for Zero Example Video Retrieval*
- Studied and implemented a contrastive loss based on MIL-NCE for the system

Video Temporal Labeling Tool

- A side project developed during my working as a research intern at McGill CIM
- GUI-supported video labeling tool implemented in Python3 with OpenCV
- Available on GitHub

AI Player for Saboteur

- A group project from COMP424 at McGill, which a class competition of playing a two-player board game Saboteur
- Implemented with annealing algorithm
- Achieved a rank of 7/161 in a tournament competition (Top 5% in class)

Modified MNIST

- A group project from COMP551 at McGill, which is a class competition on a modified MNIST dataset (MNIST with noise added to background)
- Developed our model based on LeNet 5
- Achieved an accuracy of 0.97466