HAO WEN

+86(173)2833-3196 ◇ Guangzhou, China

wenhao2018@hotmail.com ◇ https://www.linkedin.com/in/hao-wen-116911187/ ◇ https://www.richardwwhh.top/

OBJECTIVE

I am a graduated Master student from University of Florida, and I am actively seeking for a **Ph.D.** or remote **Research Assistant** position in AI and Internet of Things (IoT) related domain.

EDUCATION

M.Sc. in Electrical Computer Engineering, **University of Florida**   Sep. 2019 - Dec. 2021
Courses: Pattern Recognition, Advanced Data Structure, Image Processing & Computer Vision, and IoT Design.
GPA: 3.63/4.0

B.Eng. in Artificial Intelligence & Technology, **South China University of Technology**   Sep. 2015 - Jul. 2019
GPA: 3.26/4.0
Thesis: Electronic equipment failure detection and forecast analysis using SVM and LSTM

WORK EXPERIENCE

Machine Learning and Software Back-end Engineer   Jun. 2022 - Jun. 2023
China Merchants Bank (CMB)   *Shenzhen, China*
Designed a real-time multi-task system to detect different applications' reliability and availability in millisecond-wise with HMM.

- Managed massive growth of semi-structural data from "CMB mobile" (招商银行 APP) (An application owned by CMB with 100 million users) by using **MongoDB** APIs within **Java**.
- Segmented serialized data into fine-grained metrics to evaluate reliability and availability of "CMB mobile" using Hidden Markov Model (HMM), increased the precision of detecting vulnerable system by 3%.
- Developed a data **Retrieve and Visualization** service website based on **RESTful** APIs with **SpringBoot**.

RESEARCH EXPERIENCE

Research Intern   Apr. 2021 - Aug. 2021
STCA NLP Group   *Chinese Academy of Sciences, China*

- Improved a bidirectional Transformer based model, and further applied it for a Machine translation task with 1 BLEU score improved.
- Improved a knowledge-aware Dialogue Generation task based on Questioning Answering, applied it for multi-modal tasks with an overall improved performance.

PUBLICATIONS

- Xiao Hu[1], Hao Wen[1], "Research on Model Compression for Embedded Platform through Quantization and Pruning" (ICAITA) [Paper] [Code]

COURSE PROJECTS & COMPETITION

Model Compression Related Work   Nov. 2020 - Apr. 2021

- Constructed a CNN model on the embedded platform for image recognition, **quantized** the model to reduce overall running (Inference + Evalution) time on GPU while maintaining accuracy.
- Rebuilt a model based on Yolo-v3 for video recognition through Pruning, and gained a faster running speed on GPU without reducing accuracy.
- Achieved the training time of the quantified AlexNet model reduced by 30% amid accuracy declined only 0.08.

- Improved the precision of Yolo-v3 model after pruning, achieved the recognition of each frame reduced by about 2 latency FPS and reached a more accurate recognition of the objects in a video than the original model.

- Collabrated in writing a paper and published on **ICAITA 2021**.

Kaggle Competition on Severity of Comments   Nov. 2021 - Jan. 2022
[Code]

- Compared a pretrained RoBERTa model with TFIDF+Ridge Regression Model on various level of data cleansing, and proved RoBERTa's efficiency and accuracy on extracting comprehensive emotions from a single comment.

- Evaluated different pretrained BERT-like model for Upstream Task: RoBERTa, DeBERTa, Electra, then combined a CNN for downstream tasks to fine-tune, finally reached a 0.80893 accuracy on LeaderBoard.

TECHNICAL SKILLS

Libraries & Tools	Git, LaTeX, MySQL, MongoDB
Frameworks	Pytorch, mmcv, opencv2, SpringBoot
Languages	Java, Python, SQL, Shell Script, HTML5/CSS, MATLAB

AWARDS

- Kaggle "Jigsaw Rate Toxic Severity of Comments": **Silver Medal, (Top 2%, 18/2301)**.

- The 6th China International College Students "Internet+" Innovation and Entrepreneurship Competition, Silver Medal: **(Top 3%)**