

Han MENG

+86 15810233551 | han.meng@u.nus.edu

EDUCATION

National University of Singapore (NUS)

Ph.D. in Computer Science, School of Computing

Singapore

08/2023-Present

Beijing University of Posts and Telecommunications (BUPT)

B. Eng in Software Engineering, School of Computer Science, Minor in Intelligent Robotics

Beijing, China

09/2019-07/2023

- **GPA: 91.75; Ranking: 1/161**

- Studied in the Talent Education Experimental Class at Beijing No. 8 High School

University of Cambridge

Summer School, St. John College

Cambridge, UK

07/2021-09/2021

- Selected Course: Deep Reinforcement Learning (A)

PUBLICATIONS

- **Paper:** Han Meng, Xiaosong He, Zexing Chen, Feng Zhou, [IFDID: Information Filter upon Diversity-Improved Decoding for Diversity-Faithfulness Trade-off in NLG](#), 2023 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2023) - Rejected.
- **Paper:** Han Meng, [Smart Continuous Glucose Monitoring Prediction with Grid Long Short-Term Memory](#), Proceedings of the 7th International Conference on Cyber Security and Information Engineering (ICCSIE 2022) - Accepted.
- **Patent for Invention:** Ziwang Zhao, Han Meng, Linmei Hu, Application of Multimodal Fusion Technology Based on Mutual Learning Network in Fake News Detection (MMNet) - Published.
- **Patent for Invention:** Xiaosong He, Han Meng, A Decoding Strategy that Balances Diversity and Faithfulness - In progress.
- **Software Copyright:** Han Meng, Qi Liu, Zixuan Wei, Yang Yu, Danyang Pei, Sihai Wang, "Tang'an" - Android-based Blood Glucose Management Platform - Published.

RESEARCH INTERNSHIP EXPERIENCES

OPPO Beijing Research Institute, Department of Speech & NLP Research, Tensor Lab

06/2022-02/2023

Research Intern, Mentor: Xiaosong He

- **Research Project 1: Information Filter upon Diversity-Improved Decoding for Diversity-Faithfulness Trade-off in NLG**
 - **Background:** most of the tasks in **natural language generation (NLG)** require both diversity and faithfulness, while the traditional strategies generate repetitive texts. Guided decoding strategies proposed to enhance diversity may be unfaithful. The balance between diversity and unfaithfulness is an urgent issue to be solved.
 - **Preparation work:** mastered the knowledge of pre-trained model (PTM) architecture and training strategies of **T5, MASS, GPT2, UniLM, SimBERT, RoFormer, ELECTRA, BART**. Added cache mechanism to T5 model to speed up inference.
 - **Innovation point:** proposed an innovative decoding strategy named IFDID which can flexibly balance the diversity and faithfulness, and another novel decoding strategy named IFDID-SIMI used to increase the diversity based on a modest reduction of the faithfulness, and theoretically discussed the influencing factors of degeneration.
 - **Method:** IFDID is a two-stage decoding strategy leveraging Enhance-Filter framework, for which achieves the trade-off by increasing the probability of some typical tokens and subsequently filters them by their information amount.
 - **Verification:** designed experiments to compare the performances of benchmarks on story generation, commonsense reasoning, specific style text generation tasks. The proposed approach achieves 1.24 higher ROUGE score describing faithfulness as well as higher diversity represented by 62.5% higher upon Dist-2 than traditional approaches.
 - **Result:** completed a paper titled [IFDID: Information Filter upon Diversity-Improved Decoding for Diversity-Faithfulness Trade-off in NLG](#) as above in three months.
- **Research Project 2: Controllable Paraphrase Generation Keeping Specific Words Unchanged with Syntactic Guidance**
 - **Background:** in the paraphrase generation task, source sentences often contain phrases that should not be changed, like person's name, etc. These phrases, however, can be context-dependent and can vary by application.
 - **Motivation:** recent work has started exploring the incorporation of complex syntactic-guidance as constraints.
 - **My work:** fine-tuned a BART model without any restrict and with tagging (put two tags around the words that should not be changed) as two baselines, investigated the dataset thoroughly. The experimental datasets are MSCOCO and Quora.
 - **Next-step work:** reproduce some existing papers (like CRL-EGPG) about syntax-guided controllable paraphrase generation.

China Telecom Research Institute, Knowledge Graph Dialogue Research Group

03/2022-07/2022

Research Intern, Mentor: Peixia Sun, Zipeng Chen

- Completed the research report on the field of **task-oriented dialogue (TOD) system and dialogue state tracking (DST)** and gained a general understanding of the composition of dialogue systems (NLU, DM, NLG), and the basic concepts of dialogue state tracking (such as slot, slot group, topic switching, and state switching, etc.).
- Reproduced some important baselines (such as SOM-DST); learned about the format of the input data, evaluation indicators and methods, and frequently-used models in the process of reproducing the paper.

RESEARCH & PROJECT EXPERIENCES

Natural Language Processing (NLP) in Information Extraction (IE), LIKE LAB, BUPT

11/2021-11/2022

Research Assistant to Professor Linmei Hu

- **Background:** knowledge graph (KG) plays a decisive role in extracting named entity, relationship, event and attribute for text recognition.
- **My work:** 1) accomplished the **Named Entity Recognition (NER)**, utilized **BERT pre-trained model (PTM)** to fine-tune the downstream tasks on a given text dataset, which reached a F1 score **85.3%**.
- **My work:** 2) completed information extraction of a given text and constructed a KG; visualized the KG and batch-imported nodes and relationships based on Neo4j. In the KG, entity nodes and their images are linked.

- **My work:** 3) extracted events and relationships according to entity extraction results from other group members, sorted out relationship trigger words, distilled triples of relationships, events, and entity attributes using a rule-based method, and finally reverted to the original text.
- **Result:** the recognition accuracy for entity extraction is 85.3%; the recognition accuracies for extracting relationship, event and attribute are all over 85%; built a large-scale mapping knowledge domain including more than 3000 nodes and deployed to the neo4j platform.

NLP in Abstractive Summarization, Computer Network Information Center, Chinese Academy of Sciences 10/2021-04/2022
Research Assistant to Professor Haibo Wu

- Investigated the summarization generation algorithms and methods, familiarized with the principles and implementation methods of **Recurrent Neural Network (RNN)**, **Convolutional Neural Network (CNN)**, **LSTM**, **Attention Mechanism**, **Transformer**, **Generative Adversarial Network (GAN)**, **BERT (pre-trained model)**, **GPT (pre-trained model)** etc.
- Intensively read about 30 review papers on **automatic abstractive summarization**, then investigated the domestic and foreign cases, related principles, and methods around three research topics -- summarization optimization based on factual error correction; candidate summarization generation based on semantic similarity mask, modified sentence reordering and sentence shuffling; summarization improvement based on the global external KG; finally finished NFSC research report.

Undergraduate Innovation and Entrepreneurship Project, Project Leader

08/2020-07/2022

"Tang'an" - Android-based Blood Glucose Manager

- Aimed at connecting the non-invasive blood glucose measurement technology based on near-infrared spectroscopy to the Android APP via "Tang'an" Bluetooth module, and integrating algorithms for the prediction and processing of blood glucose.
- Responsible for blood glucose calibration and fitting, utilizing machine learning (ML) to complete the correction algorithm of error analysis designing the interface of **"Tang'an" APP** and modeled 3D shell of non-invasive blood glucose meter at the hardware.
- Won the 1) second prize of Imagine Cup, 2) third prize of Challenge Cup, 3) second prize of "Internet+" Competition, 4) national-level gold award of College Students' Innovation Competition and obtained the **software copyright**.

Cambridge University Research Training Program - Online Research Program in Artificial Intelligence 01/2022-02/2022

- Learned the writing method of the research proposal; learned about the principles and applications of **graph neural networks (GNN)**, **Transformers**, and **CNN**, **R-CNN**, **Attention model**, **BERT**, **Encoder- Decoder model**, **GAN**, etc.
- Inspired by the "Tang'an" APP and applied a **grid LSTM algorithm** to predict the blood glucose trend of the diabetic patient and **submitted a paper** *Smart Continuous Glucose Monitoring Prediction with Grid Long Short-Term Memory* to ICCSIE. Also, this is used in "Tang'an" APP.

Nanyang Technological University - Python for Data Analytics

01/2022-02/2022

- Acquired the knowledge of Python and data processing, applied **linear regression**, **logistic regression**, **support vector machine (SVM)**, **k-NN** to process data, and finally introduced several evaluation methods and regularization concepts.

2021 UCLA "PBL Interdisciplinary Online Course Project" - Big Data and Public Health

06/2021-09/2021

- Applied decision tree and **back propagation (BP) models** to the field of public health to complete electronic health records, causal reasoning, assess the medical model and then predict the spreading trend of epidemic disease.
- Responsible for writing the research report titled *The Invasion of People's Privacy in the Public Health Field in the Context of Big Data*, and finally completed the oral presentation.

2021 Innovation & Practice Program for the Cultivation of High-level Internationalized Talents of PBL Interdisciplinary Project - AI & Deep Reinforcement Learning

06/2021-09/2021

- Completed the draft of a paper entitled *Facial Keypoint Detection-CNN + Aug + TL*, and introduced a face recognition model combining neural network, augmentation, and transfer learning, with the recognition accuracy **93.4%**.

2021 MIT "Machine Learning Plus" in Business Analytics

12/2020-02/2021

- Learned the concepts and principles of perceptron, linear regression, nonlinear feature and kernel, neural network, unsupervised learning EM+ clustering, and recommendation system.
- Responsible for using Python to implement the **ridge regression model** to determine the proportion of invested stocks and writing the final report; the final model accuracy rate was **93%** which is the best result from all the members.

2021 Tunisia "the Belt and Road" Global Digital Online Study, Project Leader

12/2020-02/2021

- Learned about IoT technologies, network physics systems and robotics, mastered Bluetooth LowEnergy, LoRaWAN, MQTT Protocol, NodeRed, Fast RCNN, RCNN, YOLO, etc., and then designed innovative products.
- Applied UAV, sensor, surveillance camera, **cloud computing**, gateway, Web technology, **IoT**, target detection, and other technologies, managed to design a fully functional intelligent farm system, which made up for the closure problem of the current farm system (no malleability, no more interaction with external enterprises), and wrote a summary report.

AWARDS AND HONORS

- | | |
|---|------------------|
| • National Scholarship (Top 0.5%) | 10/2020, 10/2021 |
| • University-level Merit Students (Top 3%) | 10/2021, 10/2022 |
| • University-level First-class Academic Scholarship (Top 2%) | 10/2022 |
| • The 15 th BUPT Program Design Competition Bronze Award (Top 10%) | 10/2021 |
| • University-level Outstanding Student Cadre (Top 5%) | 10/2020 |
| • Pioneering Student Model of the School of Computer Science (Top 1%) | 01/2021, 12/2021 |
| • Second prize of the 7th China International "Internet+" College Student Innovation and Entrepreneurship Project | 06/2021 |
| • Second prize of the 17th "Challenge Cup" In-Class Academic and Technological Works Competition | 04/2021 |
| • Second prize of 2021 Microsoft "Imagine Cup" Global Student Science and Technology Competition | 12/2020 |
| • National-level gold award of College Students' Innovation and Entrepreneurship Competition | 05/2022 |

SKILLS AND INTERESTS

Computer skills: Good Command of C++, Python, C, Java SE, PyTorch; Familiar with Matlab, SQL, Axure, R, SpringBoot.

Interests: Documentaries, Reading, Rock climbing.