

HETONG WANG

+44 7715097314 ◇ s2308470@ed.ac.uk ◇ <https://erikaawang.github.io/>

EDUCATION

University of Edinburgh, UK
MSc in Artificial Intelligence

Sep 2022 – Nov 2023
GPA: 3.9/4.0

- *Relavant courses:* Machine Learning and Pattern Recognition (Theory&Practical), Natural Language Processing (Foundation&Advanced), Probabilistic Modelling and Reasoning

University of Liverpool, UK
BSc (Hons) in Computer Science

Sep 2018 – Jul 2022
GPA: 4.0 (top 1%)

- *Relavant courses:* Artificial Intelligence(Foundation&Advanced), Complexity of Algorithms, Calculus, Linear Algebra
- Exchange at Xi'an Jiaotong-Liverpool University, China

HONORS & AWARDS

British Computer Society(BCS) Prize: *top 1* student with excellent academic performance in Department of Computer Science, the University of Liverpool

University Academic Excellence Award: appr. £5,400 for each academic year, only awarded to top 1% students across the University

RESEARCH EXPERIENCE

Understanding the dynamics of implicit domain alignment in LLMs
Master Thesis in the University of Edinburgh, Supervisor: Edoardo Ponti.

May 2023 - Present

- Hypothesis: Zero-shot transfer capabilities are predicated on the ability of language models to implicitly align different domains even without parallel data, e.g. different languages will activate the same sub-networks of a multilingual LLM during inference
- Method: use Intrinsic Probing, which is a variational Bayesian model to reason the set of neurons that most correlate with a specific linguistic feature; Inspect the trend of emergence by comparing the overlap of that neuron sets of different languages in BLOOM throughout pretraining
- Correlation Analysis: We first measure the zero-shot cross-lingual transfer capability of BLOOM by XTREME benchmark; Then, we observed a great correlation between the neuron overlap rate and the zero-shot transfer performance, which confirms our hypothesis
- Github: <https://github.com/ErikaaWang/multilingual-typology-probing>

Sparse Finetuning for Instruction-tuned Language Model
Course Project in Machine Learning Practical, University of Edinburgh

Jan 2023 - May 2023

- Proposed a parameter-efficient Sparse FineTuning(SFT) method on instruction-tuned language model, which could avoid the catastrophic forgetting problem in traditional fine-tuning method and improve zero-shot performance(an improvement of 7.4 in ROUGE score on unseen data).
- Based on the modularity of Deep Neural Network(DNN): DNN could be disentangled to different facet that corresponds to a specific task ability, e.g. two different-sourced datasets of Question Generation task will modify the same subset of parameters while fine-tuning, which could be masked fine-tuned to improve the generalisation performance.

Reinforcement Learning for Robot Laboratory Skills

Jun 2022 – Sep 2022

Research Assistant in Leverhulme Research Centre, Liverpool, Supervisor: Gabriella Pizzuto.

- Paper: <https://arxiv.org/abs/2209.14875>
- This project aimed at researching and developing a robot learning framework to accelerate lab automation using deep reinforcement learning for the acquisition of new contact-rich skills
- Modified panda-gym, a simulation environment for robotic learning based on the Physical engine Pybullet and OpenAI gym to fit our laboratory environment.
- Experimented with different deep reinforcement learning algorithms(DDPG, SAC, TQC+HER) and visualised their performances to improve our reward function.
- Applied the learning framework of the simulation to the real robot task. Designed a whole experiment and evaluated the trained model by its generalization performance.

Reinforcement Learning: AI helper in Yahtzee Game

Sep 2021 – May 2022

Bachelor Thesis in the University of Liverpool, Supervisor: Michele Zito.

- This project aimed at developing a software embedded with an AI helper, which can analyse the current game state and provide action suggestions to users.
- Gained a global understanding of reinforcement learning principles, such as the Markov decision process, Monte Carlo Method. Familiar with commonly used algorithms and their principles, such as Q-learning, Sarsa, Deep Q Network.
- Developed a Yahtzee reinforcement learning environment following OpenAI gym interface, built a Deep Q Network model with Python library TensorFlow. Handled data transferring and processing, also designed a simplified state representing scheme to reduce computational cost.

SKILLS

Research Skills

- Deep Learning Toolkits: Huggingface, PyTorch, Numpy, Keras, Scikit-learn, OpenAI gym
- Paper Writing: LaTeX, Academic English Writing and Presenting
- HPC Cluster: Slurm, SGE(Sun Grid Engine), Shell, Linux

Programming Languages Python, Java, R, C/C++, SQL