Wenbo Wang

Tel: (86)13585019935 | Email: 2021211028@stu.hit.edu.cn

EDUCATION

Harbin Institute of Technology (University of Project 985) **Shandong, China**

Bachelor's Degree in Artificial Intelligence 2021.09 – Expected 2025.06

- **GPA**: 88.93/100
- **Language**: *IELTS* 6.5 | *GRE* 321(verbal:153 quantitative:168) +3 | *CET6* 587
- **Programming**: Python, Matlab, Latex
- **Core Course**: Deep Learning, AI Computing Systems, Computer Vision, Natural Language Processing, Speech Signal Processing, Circuits and Electronics, Fundamentals of signal processing, etc.

PAPER&PATENT

- Xu, H.,Wang, Z. Yan, Y.,Dong, Z., He, C.,**Wang W**., … & Xu, Y. MDReBase-KGI: A KG-Augmented database for microbe-disease relationships. Computational and Structural Biotechnology Journal (IF:4.4) **in press**
- WenboWang.2023. Intestinal microbial biota relationship extraction system based on crawlers and deep learning.CN.Patent Application 2023SR0492166.Patent Pending.

RESEARCH EXPERIENCE

Predicting the relationship between microbes and diseases in human **Harbin Institute of Technology, Weihai**

Core Member 2022.10-2023.10

Overview: Extract disease and microbe entities and the relationships between them from massive literature, then train a modified GAT model for relationship prediction by the processed data.

- Extract disease and microbe entities and sentences from crawled texts by a *Bert-based NER* model and entity normalization to find useful medical conclusions related to their internal and mutual interactions from massive Pubmed corpus
- Establish a relation type system of three types (disease and disease, microbe and microbe, and disease and microbe) in the literature by analyzing the frequency of entity relationship types of medical conclusions in the literature.
- Train a *GAT-based* model for multi-relation prediction between disease and microbe using data from LLM-augmented MDIDB database and Pubmed corpus. Use disease and microbe similarity matrix calculated from authoritative disease classification system and microbe distribute frequencies in different organs.

fMRI analysis of cognitive learning tasks **Shanghai Jiao Tong University**

Research Assistant 2024.01-2024.3

- Preprocess fMRI task data by doing correction and smoothing using ANFI and FSLeyes.
- Register and normalize fMRI task data between anatomical template and anatomical, functional individual data
- Reconstruct cortical surface from fMRI task data using FreeSurfer.

Stiffness in nonlinear brain connectome for brain mechanisms of ASD **Hong Kong Baptist University**

Summer Research Programme Member 2024.07-2024.8

- Use linear regression of gradient descent to replace the least squares method to solve the problem of slow parameter solving for linear combinations of high-dimensional vectors
- In K-fold cross-validation, use stratified sampling based on the L2 norm of the vector to replace random sampling to improve the stability of model evaluation.
- Plot the Pearson coefficient correlation matrix to evaluate the stability of the final vector in multiple fittings

PROJECT EXPERIENCE

Multimodal fake media content verification in social networks

Core Member 2023.8-2023.10

- *BLIP(Bootstrapped Language-Image Pre-training)* and *paddle-ocr* are used to simultaneously extract the content information and text content from images. *BLIP* is a model for vision-language generation tasks with vision-language joint embedding.
- Ensure the quality of the training data by computing the relationship between the text information, image content, and the title to filter out low-quality text using *Bert-whitening*.

- Process the aggregated textual information (title and image text description) by *Ernie-m*. *Ernie-m* is a multilingual pre-trained language model based on *Bert*.

Use NA-Unet to complete fundus vascular semantic segmentation

Independent Project 2023.3

✓ Recognize fundus blood vessels from pictures of fundus by a Unet semantic segmentation model.

✓ Enhance the model's focus on important information in the input by modifying the Unet model with Normalization-based Attention layers for better recognition accuracy.

✓ Augment image training dataset by image processing techniques like geometric transformations, noise injection and padding.

AWARDS&HONOURS

2021.12 People's Scholarship of the First Class (5%) Harbin Institute of Technology,Weihai

※ In reward of study ability

2022.12 Future Technology Taihu Scholarship (3%) Harbin Institute of Technology,Weihai

※ In reward of attempts in competitions and scientific research

2023.11 2nd provincial-level in CUMCM-2023 China Society for Industrial and Applied Mathematics