Xinyuan Wang

+86 15529369695 | wxy0713@stu.xjtu.edu.cn

EDUCATION BACKGROUND

Xi'an Jiaotong University (XJTU) *Xi'an, China*

Candidate for Master of Engineering in Automation; Average Score: 84.88/100 *Expected Jun 2023*

- Concentration: Natural Language Processing (Advisor: Prof. Qinke Peng)

Xi'an Jiaotong University (XJTU) *Xi'an, China*

Bachelor of Engineering in Automation; Average Score: 88.57/100 *Jun 2020*

PUBLICATIONS

[1] Wang, Ying & Peng, Qinke & Mou, Xu & **Wang, Xinyuan** & Li, Haozhou & Han, Tian & Sun, Zhao & Wang, Xiao. (2022). *A successful hybrid deep learning model aiming at promoter identification*. BMC Bioinformatics. DOI: 23. 10.1186/s12859-022-04735-6

[2] **Wang, Xinyuan** & Peng, Qinke & Mou, Xu & Li, Haozhou & Wang, Ying. *A hierarchal BERT structure for native speaker writing detection*. (Accepted by China Automation Congress CAC2022)

EXPERIENCES

Intelligent Medical Dialogue System Development *Apr 2022 – Present*

- Collected and collated data on medical conversation websites, designed and implemented the system using Python and Raspberry Pi
- Implemented a pre-triage system using the Roberta-based text classification method, designed a medical question-answering system using GPT-2 based on the self-made medical knowledge graph, realized the function of intelligent voice broadcasting of medical triage using Raspberry Pi
- Planning to enhance the fluency and logicality of the text generated by the question-answering system using reinforcement learning

Summary-enhanced Sentiment Analysis of Financial Research Reports *May 2022 – Present*

- Extracted the abstract from the long financial reports using BART, combined abstract text and the original text as the input data of model
- Constructed DD-LSTM hierarchal structure to study the implicit sentiments of financial analysts, fine-tuned the model to achieve higher performance in sentiment identification

A Hierarchal BERT Structure for Native Speaker Writing Detection *Aug 2021 – Jun 2022*

- Gathered articles' abstracts from Web of Science (WOS) to create the dataset
- Cut long text into segments via the sliding window, then got segment representations through BERT, and finally got temporal and interaction information among segments using LSTM and GAT
- Conducted ablation experiments to test the necessity of each component of the model, verified the effectiveness of the model with various classic text classification datasets

A Hybrid Deep Learning Model for Eukaryotic Promoter Identification *Apr 2021-Jun 2021*

- Proposed a novel and effective hybrid deep learning model by combining a residual network and a convolution network to extract more information in the eukaryotic gene promoter sequence
- Applied a variety of loss functions composed of cross-entropy loss, focal loss, and center loss to improve the performance
- Verified the effect of the hybrid model in actual eukaryotic promoter identification datasets

Special Equipment Data Visualization Platform *Oct 2020 – May 2021*

- Implemented the law and regulation conflict detection algorithm based on the siamese network algorithm and LSTM network
- Summarized typical fault data based on clustering algorithms, used NER to extract entities and relations to construct a knowledge graph
- Developed a platform in HTML to realize the visualization of security management workflow and real-time data for special equipment

Intelligent Fault Diagnosis System Development Based on NLP *Jun 2020-Sep 2021*

- Designed and implemented algorithms to realize the fault location according to the language description of the fault and achieved statistics of data labels based on machine learning in Python
- Developed the interface and applied the trained model to realize the fault locating and data statistics according to the existing disordered fault statistics data in real practice

Xinyuan Wang

+86 15529369695 | wxy0713@stu.xjtu.edu.cn

Teaching AI Play Flappy Bird Using Reinforcement Learning *Apr 2021 – May 2021*

- Reproduced the original codes of the Flappy Bird game in Pygame, trained the agent to play the game using CNN-based DQN
- Reduced the training time by replacing the original pictures with human-defined data

Stock Trading Strategy Study Based on Deep Reinforcement Learning *Jan 2021 – Jan 2021*

- Described the stock market by a variety of indicators in the past 30 days, set 3 operations of buying, selling, and holding as the action, defined the yield rate as the reward to design a DQN model to study the optimal strategy in stock trading
- Gave the mathematical representation of the entire reinforcement learning algorithms, verified the algorithm in the real historical data of the Chinese stock market

The Application of Named Entity Recognition in Industrial Big Data *Nov 2020 – Apr 2021*

- Normalized the text and labeled name entities in case description texts as BIOs
- Fed the embedding of different levels in the BERT model as input to the BiLSTM to extract features, tuned parameters to improve the model performance, employed the model in accident cases of industrial
- Built a visualization platform to realize the functions of statistics and matching regarding industrial accident cases in the database

Image Retrieval Using Contourlet Transform *Oct 2019 – May 2020*

- Designed and implemented the image retrieval algorithm with MATLAB to find the most similar graphics in the image library
- Adopted Contourlet Transform to extract the low-frequency and high-frequency features of the image respectively, weighted them according to a certain proportion to obtain the image algorithm with high accuracy, and verified the algorithm in the Brodatz database
- Constructed the front-end application and wrote an academic paper

Network Traffic Feature Extraction System Based on Semantic Analysis *May 2019-May 2020*

Project Leader | Undergraduate Student Research Training and Innovative Practice Program

- Employed LDA model to extract the keyword sets in the load texts in network traffic, adopted Apriori algorithm to learn the association rules between keywords in the keyword sets and generate frequent item sets, successfully reducing the data volume by 10^2 times
- Designed and implemented the relation extraction algorithm with MATLAB, tested the performance of the algorithm in content semantic analysis with real network data

HONORS & AWARDS

The 2nd Class XJTU Scholarship *Nov* 2021
The 3rd Prize of Postgraduate Group, 2021 Asia and Pacific Mathematical Contest in Modeling *Mar 2022*
XJTU Outstanding Graduates Leader *Apr 2020*
XJTU Outstanding Student *Nov 2018*
XJTU Outstanding Student Leader *Dec 2017*
XJTU Undergrad Scholarship *2016-2018*
Outstanding Leader in XJTU Students' Association Union Dec 2016

COMPUTER SKILLS

Programming Language: Python / MATLAB / C / LaTex
Algorithm and Model: Machine Learning / Deep Reinforcement Learning / Pre-trained Language Model
Library and Framework: Pytorch /Tensorflow /Keras / OpenCV / Scikit-learn

ACTIVITIES

Key Member | XJTU Chung Ying College Basketball Team *2016-2020*

- Award: Champion of XJTU Basketball Tournament for six semesters; the 4th Prize in Central China Zone, 2017 National College Student Basketball Challenge; The Champion of Shaanxi Province in Inter-college Game

Participant | Summer Session, University of California, Berkeley *2019*
Participant | Summer International School, Hong Kong Polytechnic University *2018*
Member | XJTU Students' Association Union 2016-2018