Janan Zhu

Email: bangzhubieren18@gmail.com
GitHub | Blog

EDUCATION

Tianjin University (TJU)   Department of Intelligence and Computing (Artificial Intelligence)   2020/09 – Now
Basic Courses: Computer Architecture, Database System, Data structure and Algorithm, Computer Networks, etc.
Professional Courses: Machine learning, Deep learning and Neural network, Natural Language Processing, Image Recognition, Speech Signal Processing, Knowledge Engineering, etc.

EXPERIENCE

User dialogue and application interaction system design   2023/06 – Now

Using large language models and a variety of machine learning and deep learning algorithms to make operating system more intelligent and convenient. Mainly responsible for dialogue robot design (large language model fine-tuning, instruction dataset construction, controllable text generation, dialogue management) and user service recommendation system construction based on knowledge graph.

Multimodal CoT dataset Construction and large-scale inference models training   2023/04 – 2023/06

Addressing the shortcomings of existing multimodal models in handling reasoning tasks and generating descriptions. Design strategies and implement them to integrate multimodal data from various sources. Process and align multimodal data using multiple language and image processing models, construct high-quality multimodal CoT(Chain-of-Thought) datasets, and train multimodal models for inference to ultimately enhance their reasoning abilities.

History knowledge graph and question-answering system   2022/09 – 2022/12

By crawling millions of pieces of historical data related to the Red Army's Long March from the internet, semantic analysis and entity relationship extraction are performed using LTP model. The structured information is then imported into Neo4j to form a knowledge graph. A question-answering system is built using SPARQL queries.

City scene recognition and map construction based on deep learning   2021/12 – 2022/06

The detection accuracy of existing object detection models for urban landmarks is analyzed, with further improvement achieved through transfer learning; street view photo datasets are analyzed to obtain precise positioning of urban landmarks using algorithms such as triangulation.

Automotive NVH data management and analysis system   2020/12 – 2021/04

Design website platform page to facilitate automotive data exchange. Multiple data analysis algorithms are written to achieve compatibility with various datasets.

PROJECTS

Twitter of China: Crawl real-time data of Chinese Community on Twitter, analyzing the overall situation of the Chinese Twitter community, detecting real-time dynamics and predicting hot topics, understanding overseas chinese communities' activity.

News Aggregator: A news aggregation tool that uses crawlers and multiple news data sources (API, RSS) to discover daily news hotspots and integrate events using data analysis tools and natural language processing methods. The length of the daily news briefing and the direction of hotspots such as AI or economics can be customized.

TagTagTag: A personal knowledge management tool. A concise tool for storing and querying knowledge, while also supporting to-do list management and password management.

SKILLS&TOOLS

Programming Languages: Python, C++, Bash, JavaScript, HTML, etc.

Databases: Neo4j, MySQL, Sqlite3, etc.

ML&DL Framework: PyTorch, TensorFlow, Scikit-Learn, Numpy, etc.

Other Tools: Docker, Git, Vim, Latex, etc.

Other Skills: Linux, web crawler, proficiency in Chinese and English.

AWARDS

First Prize in Tianjin University - Wenge Intelligent Computing Laboratory

Beiyang Self-Improvement Award (Top 1% school wide)