Yuanjiang WU

Email: 18810127655@163.com| Tel.: (86) 18810127655

EDUCATION

Beijing University of Posts and Telecommunications (BUPT) ***Beijing, CN***

- **Major 1**: Computer Science and Technology (84.91/100) *07.2023-present*
- **Major 2**: E-Commerce Engineering with Law (83.9/100) *07.2019-06.2023*
- **Core courses**: Internet Finance (96), Data Mining (95), Data Structures (92), Foundations of Big Data Technology(92), Digital Circuit and Logic Design (91), C++ (91), Introduction to Blockchain Security Technology (91), Operations Research (91), Probability and Statistics (90), Discrete Mathematics (90)

RESEARCH EXPERIENCES

High-Risk User Identification Based on Large Language Model for Telecom | *leader* *06.2024-09.2024*

- Conducted data preprocessing and feature optimization using matrix analysis and secondary association, successfully improving prediction accuracy by 8%, enhancing the high-risk user identification capability.
- Reproduced and optimized China Telecom's 7B TeleAI large language model, integrating Flash Attention and CUDA support based on PyTorch framework, utilizing an NVIDIA 3090 GPU for efficient inference, improving computational resource utilization.
- Designed an ensemble model based on large language models, integrating them with traditional machine learning models and applying the Chain of Thought method to improve logical reasoning. Compared to traditional models, the accuracy increased by 5%, with an F1 score reaching 81%.

Optimized Multimodal Model for Extracting Visual Information from Scanned PDFs *12.2023-07.2024*

- Refined a pretrained multimodal model that integrates text and image modalities to enhance feature extraction for visual information extraction. Designed a Spatial-Context-Adaptive Enhancer to improve the model's attention to spatial information, and a pointer decoder to extract relationships between recognized semantic entities from scanned PDF document images
- Integrating RC-RoPE and Soft Layout Mask mechanisms in self-attention to explicitly introduce 2D relative position information and adaptive spatial constraints, improving the performance in multi-line text and complex layout scenarios.
- Tested the model on the XFUND multilingual dataset, achieving an accuracy of 92.06%, significantly improving the accuracy and efficiency of document information extraction.

Differentiable Force Field Modeling Based on Deep Learning *09.2022-06.2023*

- Optimized and implemented a deep learning-based differentiable force field molecular dynamics model, utilizing OpenMP with 8-core processors to accelerate computations, reducing MD simulation time from 137.3 seconds to 36.2 seconds, achieving a 3.6x speedup.
- Improved the side-chain angle prediction model for proteins by optimizing the loss function, enhancing the model's accuracy in angle prediction and improving the precision of molecular simulation.
- Integrated deep learning models to optimize force field parameters and further accelerated the simulation process through GPU acceleration, providing more accurate dynamic data for subsequent protein motion trajectory simulations.

Java-Based Smart Home System | *leader* *05.2021-09.2021*

- Designed and implemented a RESTful architecture smart home control system using Java, supporting user registration, device binding, and remote control, improving system scalability. The front and back ends are separated, and AJAX is used to enhance page response speed and operational fluency.
- Developed and optimized the front-end interface using HTML, CSS, and JavaScript to enable real-time monitoring and control of devices, integrating dynamic data interaction to enhance the user experience.
- Conducted performance testing using JMeter, optimizing database query logic, reducing response time by 20%, and significantly improving system stability under high concurrency.

INTERNSHIP EXPERIENCES

Statistical Analysis of Chinese Living Standards Based on CGSS *11.2023 – present*

Peking University, Zhong Weiguo Research Group | Data Analyst *Beijing, CN*

- Utilized Python for data cleaning, merging, and preprocessing across multiple datasets, including removing duplicates, formatting variables, and handling invalid data. Visualized provincial GDP trends to ensure data accuracy and integrity.
- Built a multiple linear regression model to analyze the impact of income, education, and residence on national happiness, improving model prediction accuracy by 8%, providing a solid data foundation for decision-making.

Smart Service App Development *08.2022-09.2022*

ZTE Corporation *Beijing, CN*

- Led user needs assessment and market analysis, utilizing surveys and in-depth interviews to identify core needs of people with disabilities and shortcomings in existing apps, driving data-driven product design and user experience optimization.
- Applied Python web scraping for data collection and analysis to support precise product development, proposed differentiated design solutions, and developed an integrated platform for living, learning, and social interaction, significantly enhancing market competitiveness.

ACTIVITIES AND LEADERSHIP

ARnimal: An Educational AR App for Endangered Animal Conservation

Project Leader *06.2021 –05.2022*

- Developed an innovative and educational app based on market analysis and user research, allowing users to interact with virtual animals, raising awareness of animal conservation, and successfully increasing user engagement.
- Built an AR interactive environment using Unity3D and Vuforia, enabling 1:1 scale 3D models of endangered animals to enhance user immersion.
- Collected and analyzed data using Python and programmed the interaction logic in C#, ensuring users could interact with the AR models, improving system interactivity and usability.

HONORS AND AWARDS

- Academic Scholarship, BUPT *12.2022*
- Academic Scholarship, BUPT *12.2021*

SKILLS

- **Skills**: Python(PyTorch. Pandas, NumPy, Matplotlib), C/C++, Java, SQL, HTML/CSS/JavaScript
- **Languages**: Chinese (native), English (proficient, IELTS Band 7)