Weibing Wang

+1 (608) 609-4109 | wwang652@wisc.edu ∧ weibingwangwe@outlook.com | weibing-wang-w

Education

University of Wisconsin–Madison September 2021 - Present

Bachelor of Science, Dual Major in Computer Science and Mathematics *Madison, Wisconsin, USA*

GPA: 3.71/4.0 Credits: 156.

Honors: Dean's List (Fall 2021, Fall 2022, Fall 2023, Spring 2024), ACM (Association for Computing Machinery) Member, Golden Key International Honour Society Member.

Computer Science Courses: Data Structures and Algorithms, Artificial Intelligence, Database Management Systems, Machine Learning, UI/UX, Data Modeling, Distributed Systems, Big Data Systems.

Mathematics Courses: Linear Algebra, Calculus, Ordinary Differential Equations, Partial Differential Equations, Optimization Theory, Probability Theory, Combinatorics, Matrix in DS, Stochastic Processes, Data-Driven Systems.

Academic Experience

Research Interests: Focused on machine learning, particularly reinforcement learning, NLP, and IoT-driven optimization, with a strong interest in developing scalable, real-time systems for human-computer interaction (HCI) and large-scale logistics networks. Open to exploring the intersection of AI and real-time decision-making systems.

Researcher September 2024 - Present

Deep Learning-Driven Real-Time Behavior Recognition at JINGDONG Logistics X-Lab *Beijing, China (Remote)*

* Engineered and deployed a Transformer-based deep learning model optimized for real-time behavior recognition of couriers using sensor data from 400,000 devices. Applied model pruning techniques to reduce model size by 30%, maintaining 98% accuracy and lowering energy consumption by 25% in IoT environments.
* Implemented a multi-device IoT sensing system, integrating couriers' smartphones, PDAs, and external sensors, including color sensors for environment perception. By fusing data from accelerometers, gyroscopes, and GPS, the system improved recognition accuracy and reduced false positives by 20%.
* Led optimization of the Transformer model through pruning redundant layers, achieving a 15% reduction in inference latency, enabling real-time decision-making in courier operations for edge computing in large-scale logistics networks.

Research Assistant March 2024 - October 2024

AI Tutor Using LLMs, Advised by Dr. Meena Syamkumar *Madison, Wisconsin, USA*

* Engineered a web scraper using Selenium and TF-IDF for keyword extraction from forums and course platforms.
* Created OCR middleware with Tesseract for educational information extraction from images.
* Utilized NLTK and LangChain to implement a custom text chunking solution tailored for large-scale educational datasets. Integrated OpenAIEmbeddings for high-dimensional vectorization.
* Deployed FAISS to efficiently handle real-time, low-latency search across millions of data points. This architecture enabled scalable and precise retrieval-augmented generation (RAG) in educational tutoring systems.

Research Assistant November 2023 - September 2024

AR-Saliency, Advised by Prof. Yuhang Zhao *Madison, Wisconsin, USA*

* Implemented a cross-device spatial anchor system using Azure, improving synchronization accuracy by 20% across AR devices.
* Enhanced real-time 2D target detection capabilities by integrating both RTMDet and YOLOv8 models, achieving an 8% improvement in recognition speed and accuracy.
* Executed a Flask backend integrated with RabbitMQ for efficient task distribution, improving system responsiveness by 40%.
* Conducted model benchmarking between RTMDet and YOLOv8 to optimize object detection performance in AR environments, balancing speed and accuracy for different use cases.

Researcher July 2020 - December 2021

Data Analytics in Software Ethics, Advised by Dr. Jinan Liu *Beijing, China*

* Conducted Python analysis on developer response datasets to evaluate ethical practices.
* Engineered machine learning models, including logistic regression and random forests, for predictive user behavior analytics.
* Visualized cross-cultural disparities in ethical coding practices using Tableau and Power BI.

Work Experience

Algorithm Engineer Intern May 2024 - August 2024

JINGDONG Logistics X-Lab *Beijing, China*

* Analyzed over 5 million data records using Apache Spark to develop large-scale distributed data processing pipelines. Optimized ETL processes, increasing data ingestion speed by 30%, and implemented real-time data streaming using Spark Streaming.
* Conducted feature engineering, including feature selection, extraction, and transformation, improving model accuracy from 78% to 90%, significantly enhancing the model's reliability for real-time decision-making in logistics operations.
* Utilized transformer models for deep learning tasks (text classification, sequence prediction) and optimized models using PyTorch and hyperparameter tuning, reducing training time by 15%.

Undergraduate Teaching Assistant/Peer Mentor May 2024 - Present

Department of Computer Sciences, University of Wisconsin–Madison *Madison, Wisconsin, USA*

* Mentored 60+ students in CS571 (Building User Interfaces), co-taught by Prof. Yuhang Zhao and Cole Nelson, on topics such as human-computer interaction (HCI), web development, React, and React Native frameworks.
* Provided in-depth guidance on HCI principles, UI/UX design best practices, and accessibility standards, resulting in a 15% improvement in project usability scores for mentored students.
* Led code reviews and debugging sessions for student projects, helping build responsive and functional user interfaces using Dialogflow and Wit.AI.

Projects

Holos VR Medical System | *Python, OpenCV, AWS Lambda, Unity* Holos Inc. | May 2023 - January 2024

* Cooperated with Holos Inc. to architect and implement 3D skeletal modeling for medical data analysis and tutoring, and deployed computer vision techniques to accurately reconstruct skeletal structures in virtual environments.
* Integrated high-throughput data transmission APIs using AWS for seamless data flow between VR systems and cloud infrastructure, ensuring real-time interaction with minimal latency.
* Built and optimized a pipeline for large-scale data processing, reducing data transmission delays by 30%, which cut the average processing time from 150ms to 105ms, allowing real-time interactions in VR-based medical simulations. This improvement significantly enhanced the accuracy and usability of medical data for healthcare professionals.

Geospatial Social Media Platform | *Android Studio, Java, Firebase* September 2023 - December 2023

* Led a cross-functional team of 4 developers in the design and deployment of a modular social media platform, coordinating tasks and ensuring timely delivery of features, including multimedia sharing and interactive discussions.
* Led the integration of geolocation services and spatiotemporal models, utilizing Google Maps API to enable real-time location-based content sharing and visualization of community activity. This resulted in a scalable platform architecture that enhanced user engagement and facilitated real-time spatial data analysis.
* Directed the implementation of scalable API interactions using Firebase for authentication, data storage, and cloud messaging, ensuring seamless real-time data syncing across user devices and maintaining platform integrity.

Deep Learning Model for Climate Prediction | *PyTorch, LSTM, ResNet* September 2022 - December 2022

* Integrated automated data preprocessing pipelines using PyTorch, ensuring clean and structured input for time series analysis, reducing preprocessing time by 40%.
* Implemented LSTM and ResNet models to tackle the challenge of time series prediction in ice cover data, optimizing the models using advanced hyperparameter tuning techniques.
* By combining the sequential memory capabilities of LSTM with the feature extraction power of ResNet, the model achieved a 15% improvement in prediction accuracy (from 70% to 85%), compared to baseline models, validated through extensive cross-validation on climate datasets.

Extracurricular Honors & Awards

Accepted Paper: "Software Developers' Response to Unethical Coding Task Requests" at the 8th ICSTR, Singapore.

Under Review: "Multi-sourced AI Tutor for Non-Native English Speaker Students" at ACM SIGCSE.

Work in Progress: "Leveraging Pruned Transformer Architectures for IoT-Driven Real-Time Behavior Recognition in Large-Scale Logistics Networks" under the supervision of Prof. Desheng Zhang.