

## Educational Background

Xi'an Jiaotong-Liverpool University (XJTLU), Suzhou, China

Sept. 2022-Jul. 2026

- Degree: Bachelor's in Information and Computing Science (First Class)
- Honors: 2025 National Encouragement Scholarship, 2023-2024 XJTLU Undergraduate Academic Scholarship (Top 10%)

Imperial College London, Online: Synthetic Biology: Produce Fuels, Chemicals, and Pharmaceuticals in Microorganisms

Mar. 2025-May 2025

UC Los Angeles, Online: Electrical Engineering - IoT and UAV Array Network Communication System Design

Nov. 2024

Imperial College London, Online: Machine Learning in Biomedical Sciences and Bioengineering

Mar. 2024

UC Berkeley, Online (Self-learning): Data Structures and Algorithms

Apr. 2023

## Publication

- Wu, T., Huang, Y., & Purwanto, E. (2025). *M3RNet: Enhancing Multimodal Emotion Recognition with Memory-Augmented Transformers on Physiological, Behavioral, and Video Modalities*. IEEE Conference on Computer Vision and Pattern Recognition 2026. (Submitted)
- Wu, T., Purwanto, E., & Huang, Y. (2025). *Phy-FusionNet: A Memory-Augmented Transformer for Multimodal Emotion Recognition with Periodicity and Contextual Attention*. IEEE Transactions on Affective Computing. (Published)
- Wu, T., Huang, Y., Purwanto, E., Juwono, F., & Tang, F. (2025). *Enhancing Multilingual Emotion Classification with Attention Mechanism for Transnational Education*. In Proceedings of the 2025 International Conference on Artificial Intelligence and Education (ICAIE 2025), Suzhou, China, May 14–16, 2025. (Published)
- Oluwadahunsi, S., Dada, T., Huang, Y., Wu, T., Purwanto, E., & Sun, Y. (2025). *Reliable Machine Learning and Multi-objective Optimization Framework with Synthetic Data Augmentation for Load Capacity Prediction and Sustainable Design of Fully-encased Steel Columns*. Journal of Building Engineering. (Under Review)
- Wu, T., Huang, Y., Craig, P., & Purwanto, E. (2024). *EmoMA-Net: A Novel Model for Emotion Recognition Using Hybrid Multimodal Neural Networks in Adaptive Educational Systems*. In Proceedings of the 7<sup>th</sup> International Conference on Big Data and Education (ICBDE 2024), Trinity College, University of Oxford, UK, Sept. 24–26, 2024. (Published)

## Internship Experience

Intern of the Solution Department, Suzhou Artificial Intelligence Co., Ltd.

Aug. 2025 - Sept. 2025

- Contributed to the customization of the JeecgBoot open-source platform by fixing existing bugs and tailoring functional modules to meet business requirements; Successfully executed the platform's deployment and environment configuration (Nginx, Tomcat, MySQL, Redis, etc.), addressed dependency conflicts and developed automated scripts for one-click packaging and deployment.
- Enhanced the AI Agent functionality by optimizing the prompt engineering strategy and data processing logic and improving the model API call method, boosted response accuracy and execution efficiency; Deployed the Agent to internal servers, configured runtime parameters, and established log monitoring and issue tracking mechanisms to ensure service availability and stability.
- Established a GitLab-based CI/CD automation pipeline that encompasses code submission, automated building, testing, and deployment, improved the development and release efficiency; Prepared and maintained project technical documentation to facilitate team collaboration and knowledge transfer.

Software Development Intern (Back-end), Yangtze River Delta Information Intelligence Innovation Research Institute

May 2025 - Aug. 2025

- Undertook backend development tasks for an Industrial Internet of Things (IIoT) application, designed and implemented energy consumption dashboards of hoisting systems, air compressors, drainage pumps, and other critical equipment using the FastAPI framework; Applied Pydantic for data modeling and validation to ensure smooth data delivery and visualization.
- Constructed an energy consumption computation engine tailored to time-of-use electricity pricing policies, automated the generation of daily/monthly reports to support refined resource allocation and decision-making.
- Contributed to model service integration by deploying a trained OLS model as a RESTful API for energy consumption trend forecasting; Optimized complex SQL queries to improve system responsiveness under high concurrency scenarios.
- Carried out on-site system deployment, resolved network isolation and cache configuration to ensure stable operational rollout.
- Developed a component recognition system for automotive production lines, leveraged CLIP and DINO to enable rapid identification of industrial parts under low-data conditions.

## Research Program

Multimodal Temporal Point Process Modeling for Clinical Event Prediction (Shenzhen Research Institute of Big Data)

June 2025 - Present

- Constructed a baseline temporal point process model (TPP-T) and a multimodal fusion model (M-TPP) using the MIMIC-IV dataset, encoded and fused the mean and standard deviation of high-frequency vital signs (e.g., heart rate, respiratory rate) with event embeddings, and integrated them as continuous state vectors into an RNN core of M-TPP.
- Conducted comparative experiments on the prediction of the timing of the next hospital admission (regression) and the identification of clinical cause (classification) by leveraging patients' longitudinal admission records and time-aligned high-frequency vital signs, applied a weighted cross-entropy loss to address issues of class imbalance.
- Observed that M-TPP showed no significant performance gain over the baseline model and looked into the non-trivial challenges of multimodal fusion in clinical event prediction; Performed pattern analysis and filtered data based on feature coverage rates, suggested incorporating more clinically relevant modalities, reframing regression tasks as classification problems for clinical applicability, and enhancing the fusion architecture with M-TPP attention or M-TPP-gates for improved predictive performance.

First Author, **M3RNet / Phy-FusionNet: Enhancing Multimodal Emotion Recognition with Memory-Augmented Transformers** (Research Assistant to Dr. Erick Purwanto, XJTLU)

Sept. 2024 - Mar. 2025

**Summary:** The project advanced the multimodal emotion recognition through the design of a memory-augmented transformer architecture. The system integrates temporal memory, periodicity modeling, and efficient attention mechanisms to enhance long-term emotional feature retention and cross-modal fusion across physiological, behavioral, and video modalities.

- Introduced M3RNet, featuring the Memory Stream, an innovative component inspired by SAM2's memory mechanisms, which was specifically designed to enhance the transformer's ability to maintain long-term feature memory, crucial for capturing detailed

emotional cues in multimodal data.

- Implemented TMBL, a transformer-based multimodal binding learning framework, to fuse physiological and video data, which boosted the model's effectiveness in managing cross-modal interactions.
- Optimized and validated M3RNet's performance on multiple datasets, achieving over 9.15% improvement in accuracy on the PPB-Emo dataset and surpassing baseline models by 10.09% on the CL-Drive dataset, demonstrating its robustness and adaptability in real-world applications.
- Extended the architecture by incorporating a Fourier-based periodicity modeling module to capture intrinsic physiological rhythms and a head hybrid attention mechanism to improve attention efficiency while preserving accuracy, validated the full Phy-FusionNet model on WESAD, CL-Drive, PPB-Emo databases, achieving performance improvements of 13.75% over standard transformers and 4.86% absolute accuracy gains, supporting its applicability in real-world multimodal emotion recognition.

*First Author, EmoMA-Net: A Novel Model for Emotion Recognition Using Hybrid Multimodal Neural Networks in Adaptive Educational Systems (Advisor: Dr. Erick Purwanto, XJTU)* **July 2024**

**Summary:** The project developed the EmoMA-Net, a novel multimodal neural network for real-time emotion recognition in educational settings. By utilizing physiological data from wearable sensors, this model will enhance adaptive learning environments by providing educators with immediate feedback on student stress levels.

- Designed the multimodal neural network framework of EmoMA-Net, integrating advanced machine learning techniques such as Long Short-Term Memory (LSTM) systems and CNNs to capture temporal and spatial features of physiological data effectively.
- Incorporated the Convolutional Block Attention Module (CBAM), optimizing feature extraction by focusing on relevant features and improving model accuracy and reliability; This mechanism addressed the limitations of traditional attention systems by enhancing both channel-wise and spatial attention within the network.
- Utilized LSTM to enhance the temporal analysis capabilities of the model, allowing for better handling of time-series data and overcoming challenges such as gradient vanishing and information loss typically encountered in standard RNNs.
- Validated and tested the model using the WESAD dataset, achieving a high prediction accuracy of up to 99.66%, which demonstrated the model's superior performance and adaptability in emotion recognition tasks, highlighting its potential effectiveness in real-time, high-stakes educational environments.

### **Project Experience**

**Reliable and Optimized Machine-Learning Framework for Predicting Load Capacity of Concrete-Encased Steel (CES) Columns**

**June 2025**

- Designed and evaluated an interpretable ML framework for predicting the ultimate load-carrying capacity of CES members using a hybrid dataset (96 experimental + 300 KDE-generated samples) to improve training diversity.
- Developed and optimized seven regression models (XGBoost, CatBoost, LightGBM, Gradient Boosting, AdaBoost, Random Forest, and Linear Regression) using Optuna for hyperparameter tuning, assessed their performance using a composite scoring system that incorporates accuracy, reliability, AHP, and TOPSIS rankings metrics.
- Identified XGBoost as the optimal model; Interpreted the selected model using SHAP values and partial dependence plots to ensure alignment with structural engineering principles and enhance model transparency.

### **Daily Reading Tracker Design**

**Apr. 2025**

- Developed a secure and efficient reading tracker system with a robust user authentication module, realizing key features such as user registration, email verification, password encryption, and JWT-based session management; Refined the role assignment mechanism to enhance access control and implemented password recovery functionality.
- Enhanced the reading log management module by enabling multi-criteria search and filtering for faster data access; Modularized the back-end services with a clear hierarchical architecture and clearly defined front-end and back-end component responsibilities, utilized a security filter using Spring Security to enhance operational rights management.
- Carried out comprehensive unit and integration tests, verified system stability and performance under high concurrent user access.

### **Kaggle - RSNA 2024 Lumbar Spine Degenerative Classification**

**Sept. 2024**

- Developed a multi-stage deep learning model based on the ResNet-UNet framework designed to detect and classify degenerative lumbar spine conditions from MRI images; Utilized advanced techniques such as principal component analysis for feature extraction and dimensionality reduction, to ensure high-quality image preprocessing with methods like histogram equalization and Gaussian filtering.
- Implemented a sophisticated model structure, integrating ResNet as the encoder to extract deep features and capture multi-scale global context, and UNet for decoding with symmetrical skip connections to preserve high-resolution features and semantic context; Applied multi-task learning strategies at the classifier head to predict severity levels of various spinal conditions, which enhanced the model's diagnostic accuracy.
- Optimized classification and inference processes using weighted binary cross-entropy loss to handle class imbalance and employed temperature scaling and post-processing techniques to improve the stability of prediction probabilities; The ensemble methods and soft voting mechanisms further boosted the model's robustness during inference.

*Summer Undergraduate Research Fellow, MetaTeddy: Interactive Reality Modeling Interface for 3D Freeform Design*

**Jun. 2023-Sept. 2023**

- Developed a 2D interactive interface using Unity based on the Teddy system framework; Implemented precise cutting operations using the Computational Geometry Algorithms Library (CGAL) and partitioning algorithms, and enhanced model manipulation with the TouchScript plugin for easy rotation by users.
- Integrated Extended Reality (XR) technologies using Microsoft HoloLens 2 as the visual operating device for users; Combined MetaTeddy with AR headsets through the Unity engine and AR Foundation plugin, and used Node.js with the Socket.io library to facilitate network connections and process monitoring between the server side and AR glasses, enabling users to transform models from a 2D tablet interface into 3D visualizations.

### **Additional Information**

- **Languages:** Chinese (native), English (fluent, TOEFL 107)
- **Programming:** Python (Pandas, Numpy, Scikit-learn, Matplotlib, Seaborn, TensorFlow, PyTorch, pytest), Java (JUnit, Spring Boot), SQL (MySQL, Snowflake), C, C#, R, Matlab, Git
- **Interests:** Piano (Level 10 certification from Shanghai Conservatory of Music), Street Dance (Level 4 certification from China Street Dance Association), Badminton, Movies
- **Community Activity:** Member, Xi'an Jiaotong-Liverpool University Psychology Association (PSYA)