# JIAYI ZHANG

+86 18822185817 jiayi.zhang718@gmail.com Google Scholar GitHub Nottingham, NG7 2RD

## EDUCATION

**University of Nottingham** **Sep 2024 - Jun 2028 (Expected)**

*BSc (Hons) Mathematics with Applied Mathematics* *Nottingham, UK*

## PERSONAL INTERESTS/ PRACTICE EXPERIENCE

* For equal contribution, † Corresponding author.

[1] Z. Li*, **J. Zhang***, Y. Wang*, J. You, J. Zhong, S. Dev†, "FGSCare: A Feature-driven Grid Search-based Machine Learning Framework for Coronary Heart Disease Prediction", *Healthcare Analytics*, under review.

[2] **J. Zhang***, Y. Zhang*, Y. Zheng, Y. Wang, J. You, Y. Xu, W. Jiang, S. Dev†, "TrafficKAN-GCN: Graph Convolutional-based Kolmogorov-Arnold Network for Traffic Flow Optimization", *Decision Analytics Journal*, under review.

[3] Y. Li*, H. Wang*, **J. Zhang**, J. You, J. Xu, P. Wu, Y. Xiao, S. Dev†, "Segregation and Context Aggregation Network for Real-Time Cloud Segementation", *The International Conference on Learning Representations (ICLR) Workshop*.

[4] Z. Xu*, J. Zhong*, H. Wang, J. Xu, Y. Li, J. You, **J. Zhang**, R. Wu, S. Dev†, "RAINER: A Robust Ensemble Learning Grid Search-Tuned Framework for Rainfall Patterns Prediction", *arXiv:2501.16900*, Preprint.

## PROJECTS EXPERIENCE

**FGSCare: ML Framework for CHD Prediction** | *Machine Learning*, *Medical AI* **Mar 2025**

- Developed **FGSCare**, a feature-driven framework for **coronary heart disease prediction**, combining systematic **feature selection**, **grid search**, and **SHAP-based interpretation** to improve diagnostic accuracy and model transparency.
- Benchmarked traditional models (e.g., **Gradient Boosting, ElasticNet**) and deep learning models (e.g., **KAN, Transformer**) on the **Framingham dataset**, achieving up to **2.2% F1-score** gain with feature selection.
- Applied data preprocessing techniques (e.g., **outlier capping**, **SMOTE**, **PCA**) and identified top predictors such as **age**, **glucose**, and **blood pressure** via SHAP, enhancing clinical interpretability.

**KAN-GCN for Traffic Flow Optimization** | *Deep Learning*, *Graph Convolutional Networks* **Feb 2025**

- Developed **KAN-GCN**, a hybrid model for urban **traffic flow optimization**, integrating **Kolmogorov–Arnold Networks (KAN)** with **Graph Convolutional Networks (GCN)** to enhance non-linear traffic pattern modeling.
- Implemented an adaptive feature transformation in KAN, improving function approximations and preserving spatial dependencies in GCN. This enhances the model's ability to capture complex traffic patterns, boosting accuracy and robustness.
- Conducted extensive evaluations on real-world traffic datasets, comparing **KAN-GCN** with MLP-GCN and decision trees. Results demonstrate **robustness to noisy data** and adaptability to dynamic traffic conditions, reducing RMSE by 2.7% in complex scenarios.
- Analyzed challenges of integrating KAN into traffic networks, identifying **overfitting risks in low-data regimes** and proposing **regularization** for stability. Explored **scalability with Transformer-based architectures** for long-term forecasting.

**Segregation and Context Aggregation Network for Cloud Segmentation** | *Deep Learning*, *Computer Vision* **Jan 2025**

- Developed **SCANet**, a lightweight learning-based visual model for real-time **cloud segmentation**, integrating a novel **Segregation and Context Aggregation Module (SCAM)** to enhance cloud-sky feature separation with minimal computational cost.
- Implemented an pre-training strategy leveraging the **SWINySEG dataset**, eliminating dependency on ImageNet pre-training and improving segmentation performance. SCANet-lite achieves **1390 FPS** in FP16, surpassing real-time processing standards.
- Evaluated SCANet on **day-time, night-time, and full SWINySEG datasets**, achieving **state-of-the-art performance** with 70.9% fewer parameters compared to prior methods. SCANet-large attains **97.0% accuracy** while maintaining computational efficiency.

**RAINER: Grid-Tuned Ensemble Learning for Accurate Rainfall Prediction** | *Rainfall Prediction*, *Machine Learning* **Dec 2024**

- Developed **RAINER**, a robust ensemble learning framework for **rainfall prediction**, utilizing a **grid search-tuned** approach to optimize model selection and hyperparameters, significantly enhancing forecasting accuracy across diverse climatic conditions.
- Implemented a **heterogeneous ensemble model** that integrates **Random Forest, Gradient Boosting, XGBoost, and Support Vector Regression**, systematically fine-tuned using **grid search and cross-validation** to identify the best-performing configurations. This approach yielded superior predictive performance, outperforming individual models across multiple meteorological datasets.
- Evaluated **RAINER** on **regional meteorological datasets**, demonstrating **strong generalization** with reduced errors and **notable RMSE, MAE, and $R^2$ improvements** over traditional models across diverse conditions.
- Leveraged **feature selection and engineering** techniques to identify the most influential meteorological variables, optimizing computational efficiency while maintaining high predictive accuracy in **short-term and long-term rainfall forecasting**.

## SKILLS

**Programming Languages**: Python, Matlab, LaTeX, C, HTML, CSS, Javascript, Kotlin **Tools**: Pytorch, Tensorflow, scikit-learn, Numpy, Pandas, Matplotlib, React, Docker, Spring Boot, AWS, CUDA, Azure, Google Cloud

**Machine Learning Model**: CNN, ResNet, Faster R-CNN, YOLO, LSTM, Transformer, BERT, ViT, GANs, VAE, Stable Diffusion