YUNRUI ZHANG

+61 452560467 ◇ Sydney, NSW
yunruizhang20@gmail.com ◇ Linkedin ◇ Github

EDUCATION

Doctor of Philosophy in Computer Science, UNSW **August 2022 - Present**
Bachelor of Engineering (Honours) in Computer Engineering, UNSW **August 2018 - August 2022**

RESEARCH AREA

Time Series Analysis Time Series Classification, Time Series Forecasting
Machine Learning Improving probabilistic LLM sampling, Calibration, Prior Shift, Quantification

PUBLICATION

[1] Y Zhang, GE Batista, S Kanhere. Rethinking Time Series Classification: The Impact of Temporal Information for Classification. Accepted by Pacific-Asia Conference on Knowledge Discovery and Data Mining (PAKDD) 2025

[2] Y Zhang, GE Batista, S Kanhere. Label Shift Estimation With Incremental Prior Update. Proceedings of the 2025 SIAM International Conference on Data Mining (SDM). 2025, 134-142

[3] Y Zhang, GE Batista, S Kanhere. Instance-Wise Monotonic Calibration by Constrained Transformation. Accepted by Uncertainty in Artificial Intelligence (UAI) 2025

[4] Z Donyavi, F Li, Y Zhang, DF Silva, GE Batista Match: A Maximum-Likelihood Approach for Classification under Label Shift. Accepted by KDD 2025

[5] Y Zhang, GE Batista, S Kanhere. Mosaic: an Accurate and Efficient Kernel-based Multivariate Time Series Classifier. Under review by CIKM 2025

SKILLS

Languages Python, C++, C, VHDL, Java, SQL, Shell, and more
Tools & Frameworks LLM inference, Pytorch, Sklearn, Scipy, and more

RESEARCH PROJECTS

Improving LLM probabilistic sampling in open-end generation

- Investigated sampling strategies to improve output diversity and quality in LLM generation beyond temperature and top-k/top-p decoding.
- Introduced auxiliary information to guide token-level sampling strategies for open-ended generation tasks.
- Verified on state-of-the-art open-source LLMs of various sizes, as well as reasoning models.
- Analyzed the effect of sample entropy, diversity, and calibration under different sampling schemes, and their effects on reasoning models.

Calibration of Deep learning models

- Having a well-estimated confidence score is crucial for many safety-critical tasks, yet the probabilistic outputs of many models, especially deep learning models are often inaccurate.
- Calibration addresses the problem of miscalibrated confidence scores in modern classifiers, where predicted probabilities often fail to reflect the true posterior.
- Existing post hoc calibration methods lack expressiveness and the ability to maintain instance-wise monotonicity after calibration.
- Proposed novel calibration methods using a parameterized calibration map that ensures instance-wise monotonicity, improves interpretability, and outperforms state-of-the-art techniques.

Label Shift Estimation

- Addressed the challenge of estimating label priors under distribution shift, where test-time label proportions differ from the training set.
- Proposed a novel label shift estimation method that takes an alternative perspective by incrementally updating label priors based on individual samples.
- Achieved superior performance over state-of-the-art approaches while relying on weaker calibration, with linear time complexity and compatibility with black-box models.

Time Series Classfication

- Proposed a novel kernel-based multivariate time series classifier designed to leverage inter-dimensional correlations for improved classification performance.
- Achieved competitive results while maintaining scalability and reasonable time complexity.
- Identified a major issue in current time series classification benchmarks, where temporal information often does not contribute meaningfully to classification.
- Introduced a new benchmark variation that emphasizes the extraction and utilization of temporal information as a core evaluation criterion.

EXPERIENCE

Casual Academic Sep 2023 - Present
University of New South Wales *Sydney, NSW*

- Tutor and marker for COMP9418 Advanced Topics in Statistical Machine Learning.
- Design and organize the new format for COMP9418.

Research assistant Sep 2022 - Dec 2022
University of New South Wales *Sydney, NSW*

- Worked with Dr Sushmita Ruj and Prof Salil Kanhere on a decentralized carbon credit validation scheme.
- Collaborate with industrial partner Envirocoin on developing carbon credit validation methodologies.

Research assistant Jun 2019 - Dec 2019
University of Technology Sydney *Sydney, NSW*

- Worked with Dr Kai Wu and Prof Ghassan Beydoun on a smart farm project.
- Developed wearable devices for livestock monitoring.

AWARDS

- UNSW TFS PhD scholarship
- UNSW Taste Of Research Scholarship 2021
- Cyber Security Cooperative Research Center Honours Scholarship