Thesis Topic:

Efficient Implementation of Mondrian Forest at the Edge

Background:

In response to the demand for real-time applications such as pedestrian detection, IP address search, and nanoparticle sensors, the field of machine learning is witnessing a shift towards learning from evolving data streams. This transition blurs the boundary between training and inference. Tree-based ensembles, including random forests and gradient boosting trees, remain dominant in solving real-world data problems. However, traditional offline learning, which involves storing data prior to learning, may not be feasible for edge devices due to limited resources.

Tree-based online learning models like Hoeffding Trees and Mondrian Forest have gained attention recently. While Hoeffding Trees learn a tree model without storing data points, Mondrian Forest shows promise with competitive accuracy and predetermined tree size akin to offline models. Despite the potential efficiency optimizations demonstrated in offline-trained tree-based models, efficiently realizing Mondrian Forest remains an open challenge.

Project Description:

This project aims to delve into the intricacies of implementing Mondrian Forest efficiently. We will explore suitable design methodologies to deploy Mondrian Forest on resource-limited platforms or with FPGAs. The student will analyze the state-of-the-art techniques in the chosen research domains and devise novel methodologies to enhance efficiency fundamentally. Evaluation will include overhead analysis, prediction accuracy, and runtime comparison against baseline implementations from the state-of-the-art. A previous work by the supervisors [2] provides a reference scale for this assignment.

Key References:

1. Mondrian forests: Efficient Online Random Forests
2. Accelerated Real-Time Classification of Evolving Data Streams using Adaptive Random Forests
3. Realization of Random Forest for Real-Time Evaluation through Tree Framing
4. A Random Forest Using a Multi-valued Decision Diagram on an FPGA

Student Profile:

Skills Required:

- Knowledge of computer architecture
- Proficiency in C/HLS programming
- Basic knowledge of machine learning

Desired Knowledge/Experience:

- Advanced Computer Architecture
- Open-source software development

Supervisor/Contact Information:

- Dr. Ir. Kuan-Hsun Chen k.h.chen@utwente.nl
- Dr. Ir. Nikolaos Alachiotis n.alachiotis@utwente.nl