



Utilizing the Learnings of ML Algorithms to Depreciate Potential Difference

Apoorva Reddy Proddutoori

San Diego
apoorvaproddutoori@gmail.com

ABSTRACT

The machine learning framework is employed for power management decision-making under conditions of uncertainty, enhancing the adaptability and efficiency of the power management mechanism in response to the formulated model. Mobile software is increasingly becoming feature-rich, often integrated with the powerful decision-making capabilities of machine learning (ML). This approach is applied to complex multiprocessor-based SoCs, including CPUs, APUs, and GPUs. We address the challenge of decision-based control of various power management features in ICs, which often consist of numerous heterogeneous IPs. To tackle this, we utilize infinite-horizon fully observable MDPs to derive a policy of actions that maximizes the expected utility of the formulated Power Management utility function. This approach strikes a balance between the demand for desired performance and the goal of achieving optimal power savings, offering a significant improvement over the commonly used FSM-based power management techniques.

Keywords: multiprocessor, power management, performance, machine learning, sorting

INTRODUCTION

This paper introduces a machine learning-based reinforcement learning approach that maps Finite State Machines (FSMs), traditionally used for power management control in SoCs, to Markov Decision Process (MDP)-based agents for controlling power management features of Integrated Circuits (ICs). To meet the consequent rise in power and performance demands, system and hardware architects integrate specialized hardware units into their system-on-chips (SoCs) and frameworks for optimal compute delegation. While these SoC innovations significantly enhance ML model performance and power efficiency, the auxiliary data processing and supporting infrastructure required for ML model execution can significantly impact system performance. This paper introduces the concept of an "AI tax," referring to the time spent on non-model execution tasks. We analyze the execution pipeline of open-source machine learning benchmarks and Android applications, characterizing the ML algorithms and discussing potential difference performance bottlenecks that may arise unexpectedly.

New advances in data science and analytics are being used in many areas. For example, Machine Learning is helping to make decisions and predictions in energy management for small-scale power systems called micro-grids. These micro-grids use renewable energy, batteries, and different types of loads to provide power. As more components are added, it becomes more complex to decide how to distribute energy efficiently.

One important metric is performance per unit power. in the current chip design. Clock network frequently becomes a critical supporter (~30%) to the all out power consumption Power advancement for clock network has been studied for decades, and numerous methods like the clock dynamic voltage frequency scaling (DVFS), gating, and the size of the clock buffer etc. have already been tested in nearly all SoCs. to lessen the power supply, and the tools for implementation also support multi-bit flip- conversion using multiple regular flip-flops form a single cell when they are grouped together.

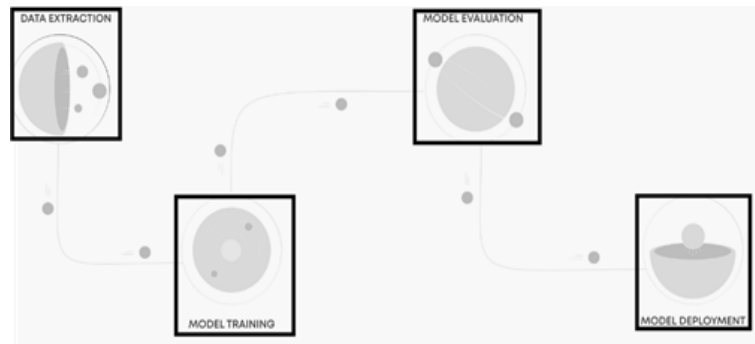


Figure 1: ML Model

This paper conducts a comparative analysis of five distinct machine learning (ML) algorithms aimed at identifying the most suitable algorithm for energy management within a community of three micro-grids. The algorithm deemed the most effective is the one that achieves the highest accuracy on the specified dataset. The analysis proceeds by creating and splitting a data set into training and testing subsets. Additionally, the importance of various features in determining the optimal energy distribution mode is established. Labels are assigned to indicate the operational modes. Section II details the implementation of seven different preprocessing methods on the data set. In Section III, the ML models are trained and applied to the test data to evaluate their accuracy. Furthermore, the visualization of the five features illustrates the classification of the labels. Section IV presents the findings of this comparison in a bar diagram, offering insights into the most effective features among the five. The final section concludes the study, summarizing the key findings and implications.

ALGORITHM MODEL

A. Data Accumulation

Gathering information from sensors might seem simple, but it can make an application's design more complex and affect how it's built. For instance, quickly taking in raw images can overload the system's memory or make it slower if the images are not clear or processing them takes too long. In some cases, multiple sensors are used, which means combining their data into one can be needed. This extra work might happen at the same time as the rest of the app, making it less efficient. Another option is to use different parts of the system, like a DSP, but this also has its own challenges, such as extra steps needed to make everything work together. We found that the code for handling data was a big reason why some apps took a long time to run.

Computational Electromagnetics offers foundational physical models for a wide range of electronic devices, including Integrated Circuits (IC), packaging, boards, and connectors. Over the years, numerous methods have been developed, with the method of moments (MoM) being a particularly popular approach used for parasitic extraction, noise analysis, signal integrity, power integrity, and more.

Machine learning (ML) is currently driving the forefront of Information Technology (IT) technology. It focuses on developing algorithms that can learn from data and make predictions, prioritizing the best predictions rather than exact solutions. This has led to its applications in scientific computations being relatively limited today. However, the vast consumer applications of machine learning have attracted significant attention, resulting in the creation of extensive computing resources, which span from software algorithms to hardware platforms.

This is what inspires us to establish a direct link between machine learning and MoM. By doing so, we aim to leverage the vast resources of machine learning to enhance scientific computing, specifically in the areas of signal integrity and power integrity modeling.

B. DNN Data

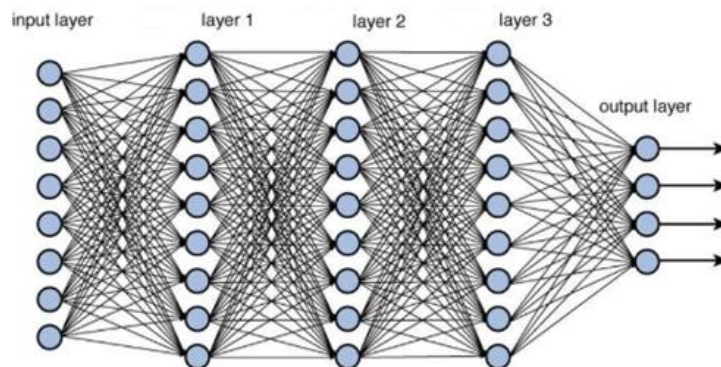


Figure 2: DNN logic

In the realm of deep neural networks, the most common training method is back-propagation. This technique adjusts the weights based on the gradient of the error function, but it requires labeled data for supervised learning. However, applying back-propagation or supervised learning to real-time, unlabeled sensory input data is challenging. To overcome this, we turn to unsupervised learning algorithms, such as spike-timing dependent plasticity (STDP), which is found in biological nervous systems. This approach is expected to enhance learning in robotic or defense applications, especially when dealing with large volumes of unlabeled data, similar to how we analyze and associate sensory input data.

In this Observation, we aim to apply the fundamental learning principles observed in biological nervous systems using advanced CMOS technology in a novel architecture. Our focus is on utilizing digital logic and memory circuits for the design of neuromorphic processors. These circuits are advantageous because they can achieve dense and low-power hardware through transistor scaling and dynamic voltage scaling, respectively. We will introduce two neuromorphic processors that incorporate plastic synapses and integrate-and-fire neurons for applications in pattern recognition and data clustering.

We will also address several challenges, including the learning capability with synaptic precision, the operation of low-voltage synapse arrays, and the integration of learning and classification functionalities within the same hardware.

MODEL PROCESSING

A. Decision Tree Model

The Decision Tree model for Economic Protection Evaluation is a machine learning hierarchical framework designed to examine various elements of economic protection development. This method leverages complex correlations between measures such as carbon emissions, GDP growth rate, and technology investment, tailoring the model to suit economic protection assessment needs. The resulting tree structure showcases a level of economic protection development at a machine learning-based economic protection development level, utilizing Decision Tree and Ensemble Algorithms trends, and aids policymakers in making informed decisions. By employing ensemble approaches such as Random Forest, the model's resilience is enhanced, overcoming the limitations of standalone Decision Trees. This hybrid approach offers a comprehensive framework for thorough analysis of economic protection development while maintaining a balance between scientific precision and practical applicability. The subsequent sections of this text delve deeper into the description of the model, attributes selection, node splitting criteria, and its application in economic protection assessment.

B. SVM Model

This paper addresses the multi-label classification problem, which is identified as an unbalanced, multi-task learning challenge. The solution proposed is a generalized version of the multi-task least-squares support vector machine (MTLS-SVM), named the multi-label least-squares support vector machine (ML^2 S-SVM). Experimental findings with emotion, scene, yeast, and BibTeX data sets demonstrate that ML^2 S-SVM outperforms current state-of-the-art methods in terms of Hamming loss and instance-based F1 score, indicating its competitive edge. The performance of ML^2 S-SVM is significantly impacted by parameter values, emphasizing the importance of careful parameter selection in advance.

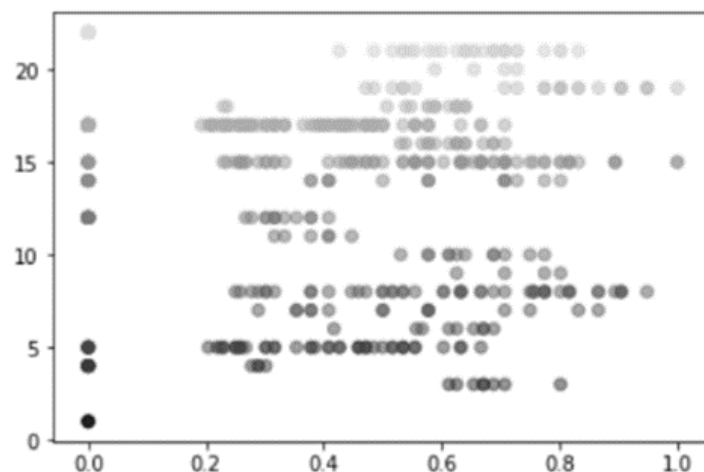


Figure 3: Temperature (C) vs Voltage (mV)

In summary, this work introduces ML^2 S-SVM, a novel approach to multi-label classification that not only addresses the imbalance issue but also uniquely models arbitrary order correlations among labels by allowing multiple labels to share a subspace.

CONCLUSION

This paper introduced a reinforcement learning-based approach for optimizing power management features in integrated circuits (ICs), particularly in complex, multi-processor-based System-on-Chips (SoCs). The methodology involves converting the existing Field-Selection Modules (FSMs) into Mathematical Decision Diagrams (MDPs) for more efficient control of power management events. An infinite-horizon, fully observable MDP was employed to develop a policy that maximizes the expected utility of a predefined utility function. This approach effectively balances the need for optimal performance with the goal of achieving the most energy-efficient power management as opposed to the traditional, static FSM-based power management techniques.

A MDP framework was utilized to simulate decision-making in power management under uncertain conditions, a critical aspect of reinforcement learning. The results highlight the potential of machine learning, especially MDP, to effectively bridge the gap between power management and performance optimization, while also accommodating changes in data and control patterns within complex SoC systems.

REFERENCES

- [1]. He Ming Yao, Yu Wei Qin, Li Jun Jiang¹, "Machine Learning Based MoM (ML-MoM) For Parasitic Capacitance Extractions", IEEE, March 2016
- [2]. David Akselrod, "ML-based Reinforcement Learning Approach for Power Management in SoCs", 2019 32nd IEEE International System-on-Chip Conference (SOCC)
- [3]. Jae-sun Seo and Mingoo Seok, "Digital CMOS Neuromorphic Processor Design Featuring Unsupervised Online Learning", IEEE, May 2015
- [4]. Michael Buch¹, Zahra Azad², Ajay Joshi², and Vijay Janapa Reddi, "AI Tax in Mobile SoCs: End-to-end Performance Analysis of Machine Learning in Smartphones", 2021 IEEE International Symposium on Performance Analysis of Systems and Software (ISPASS)
- [5]. Mahshid Khoshlessan, Babak Fahimi, Morgan Kiani, "A comparison between Machine learning algorithms for the application of micro-grids Energy management", IEEE, Feb 2020
- [6]. Santanu Kundu, Niraj A. Mehta, Arpan Sircar, "Machine Learning Based Flip-Flop Clustering for Clock Network Power Improvement", 2022 IEEE International Conference on Emerging Electronics (ICEE)
- [7]. https://www.researchgate.net/publication/352040064_Modified_ML-kNN_and_Rank_SVM_for_Multi-label_Pattern_Classification
- [8]. https://www.researchgate.net/publication/337841055_ML_2_S-SVM_multi-label_least-squares_support_vector_machine_classifiers