ABSTRACT

As machine learning systems scale in terms of data volume and model complexity, there is a growing need for methods that enhance efficiency, explainability, and robustness. We explore subset selection as a principled approach to address these challenges across diverse applications. Our objective is to identify representative elements, from data or model components, that maintain task performance, reduce resource consumption, and eliminate harmful or redundant elements that may degrade the model's performance. By focusing on core subsets within data or model structures, we aim to build efficient and resilient systems for real-world, large-scale applications.

Specifically, in the first work of this thesis, we focus on a subset-selection-based defense mechanism to enhance the security of Federated Learning (FL) systems, which are vulnerable to model poisoning attacks—particularly edge-case attacks that are difficult to detect. We propose `DataDefense`, which leverages a small external defense dataset to jointly learn a poisoned data detector and a client importance model, iteratively refining both through alternating minimization during FL rounds. This approach substantially reduces attack success rates while preserving model utility.

Our second work addresses subset selection for efficient whole-network pruning in deep convolutional neural networks (CNNs), aiming for efficient deployment on resource-constrained devices. We propose a two-level hierarchical pruning framework (`HBGTS-B`) that removes redundant filters through sparse approximation of filter weights. At the lower level, we design efficient filter selection algorithms based on orthogonal matching pursuit or backward pruning. At the higher level, we greedily select layers for pruning using strategies that minimize classification

loss. This approach achieves substantial model compression while preserving performance.

In our third work, we focus on exemplar subset selection for in-context learning (ICL) with large language models (LLMs), where performance critically depends on the quality of exemplars. We propose EXPLORA, a static exemplar subset selection method tailored to complex reasoning tasks. EXPLORA explores the exemplar space to learn a parameterized scoring function, enabling efficient evaluation and selection of exemplar subsets. This approach yields exemplars that generalize well across test examples while reducing inference latency.

Finally, our fourth contribution reformulates exemplar subset selection as a top-m best arms identification problem under a stochastic linear bandit framework. We introduce CASE (Challenger Arm Sampling for Exemplar selection), a sample-efficient algorithm that selectively explores promising exemplar subsets while maintaining a dynamic shortlist of challengers. By leveraging a linear scoring function over subsets, CASE significantly reduces the number of LLM evaluations required to identify optimal exemplars.

By addressing these challenges, this research advances the development of robust, efficient, and scalable machine learning systems through principled subset selection. The proposed methods hold broad applicability across diverse domains including federated learning, edge deployment of deep models, and large language model reasoning, paving the way for trustworthy and computationally efficient machine learning systems in real-world, resource-constrained, and adversarial environments.

Keywords: Machine Learning, Subset Selection, Federated Learning, Model Poisoning Attacks, Data-Driven Defense, Filter Pruning, In-Context Learning, Large Language Models, Stochastic Linear Bandits, Robustness, Efficiency