Project Title: Adaptive Active Learning for Regression via Reinforcement Learning
Mentor: Simon Nguyen Quarter: Fall 2025

Introduction: The Cost of Labeling In many fields, data is abundant, but labeled data is expensive. Running a single experiment to determine the properties of a chemical compound or to test a robotic control policy can be costly and time-consuming. Active Learning (AL) addresses this challenge by intelligently selecting which data points to label next rather than selecting them randomly. The goal is to train an accurate model with the fewest possible samples.

The Challenge: Exploration vs. Investigation A central challenge in active learning for regression is balancing two competing objectives. First is exploration, which involves sampling from regions of the feature space where we have no data to maximize diversity. Second is investigation, which involves sampling from regions where the model's predictions are uncertain to minimize variance.

Existing state-of-the-art methods, such as *improved Greedy Sampling (iGS)*, balance these objectives using fixed multiplicative rules. While effective in many cases, we hypothesized that static rules fail in heterogeneous environments. Specifically, they fail in "traps" where high data density conceals pockets of high model error. In these cases, a multiplicative rule mathematically suppresses the error signal, which makes the algorithm blind to the trap.

Our Approach: Weighted Improved Greedy Sampling (WiGS) To solve this, we developed a new framework called Weighted improved Greedy Sampling (WiGS). Instead of a fixed rule, WiGS uses a dynamic additive criterion governed by a weight parameter. This parameter acts as a sliding scale. It allows the algorithm to prioritize exploration or investigation based on the current needs of the model. The core innovation of our project was automating this choice using Reinforcement Learning (RL). We formulated the active learning process as a Markov Decision Process (MDP) where an agent observes the state of the dataset and selects the optimal weight for the next batch of samples. We trained a Soft Actor-Critic (SAC) agent to learn a policy that maximizes the reduction in prediction error (RMSE) over time.

Results and Conclusion We evaluated our method on two adversarial synthetic datasets and 18 real-world regression benchmarks. Our results demonstrated that the adaptive WiGS agent successfully decoupled exploration from investigation. In synthetic tests designed with high-noise "traps," the agent learned to ignore misleading feature density and focus aggressively on uncertainty reduction. This reduced error by up to 40% compared to the iGS baseline. Analysis of the learned policies showed that the agent did not converge to a single static number. Instead, it learned a dynamic strategy that prioritized exploration in complex regions and investigation in simpler noisy regions. This project validates the potential of "learning to learn" in active sampling. By allowing an AI agent to dynamically control the sampling strategy, we can achieve robust performance across diverse domains without requiring human intuition to tune hyperparameters