Hyperparameter Exploration Report

Introduction

This report outlines the hyperparameter tuning conducted for the Multi-Layer Perceptron (MLP) model designed to predict the number of bedrooms in properties listed in the New York housing market.

Hyperparameters Explored

The following hyperparameters were systematically varied during model development:

- **Number of Hidden Layers and Units Per Layer**: Varied to explore shallow versus deep network architectures.
- **Learning Rate**: Tested different scales to optimize convergence speed and stability.
- **Epochs**: Altered to find a balance between underfitting and overfitting.
- **Learning Rate Decay**: Included to investigate its impact on the learning rate adaptation over time.

Methodology

Experiments were performed using different combinations of hyperparameters to identify the configuration that yields the highest classification accuracy. The performance was primarily evaluated based on the accuracy metric.

Results

The experiments demonstrated varied successes. Key observations include:

- Networks with a higher number of hidden units and layers did not necessarily perform better, highlighting the importance of matching network capacity to the complexity of the task.
- A higher learning rate initially seemed beneficial but often led to instability in deeper networks.
- Implementing a decay in the learning rate generally allowed for more stable training across more epochs but did not always correlate with better final performance.

Best Configuration

The best-performing model configuration used two hidden layers, each with 150 units, a learning rate of 0.01, and a learning rate decay of 0.0005, trained over 500 epochs. This setup achieved the highest accuracy observed across all experiments.