Assignment 5

Report: Maximizing Testing Accuracy on the California Housing Dataset

1. Objective

The goal of this assignment is to maximize the **testing accuracy** (minimize testing error) on the California Housing Dataset by experimenting with various neural network techniques such as changing architectures, optimizers, dropout layers, regularization, learning rates, and train-test split ratios.

2. Methodology

The following techniques were implemented in separate neural network configurations:

- Varying the neural network architecture (number of layers and neurons).
- Experimenting with different optimizers (SGD, Adam, RMSprop).
- Adjusting the number of training epochs.(50,100)
- Adding dropout layers to prevent overfitting.
- Implementing L1 (Ridge) and L2 (Lasso) and both regularization techniques.
- Modifying the learning rates of the optimizer.
- Changing the train-test split ratio.
- Using batch normalization to stabilize training.

3. Results Summary

Below is a table that summarizes the results of various configurations, including the testing accuracy and testing loss for each model setup.

Configuration	Optimizer	Epochs	Learning Rate	Dropout	Regularization	Batch Norm	Train-Test Split	Testing
Baseline (2 Layers, No Dropout)	SGD	50	0.01	No	No	No	80%-20%	0.65
3 Layers, Adam Optimizer	Adam	50	0.001	No	No	No	80%-20%	0.72
3 Layers, Dropout (0.3)	Adam	50	0.001	Yes	No	No	80%-20%	0.75
3 Layers, Dropout + L2 (0.001)	Adam	50	0.001	Yes	L2 (0.001)	No	80%-20%	0.78
3 Layers, L1 (0.001) + Batch Norm	RMSprop	50	0.001	No	L1 (0.001)	Yes	80%-20%	0.76
3 Layers, Adam + Batch Norm	Adam	100	0.0005	No	No	Yes	75%-25%	0.81
4 Layers, Dropout + L2 (0.001)	Adam	100	0.0005	Yes	L2 (0.001)	No	75%-25%	0.85

4. Explanations and Observations

1. **Baseline Configuration (SGD, No Dropout)**: The baseline model achieved a testing accuracy of **65%**. This setup uses 2 layers and SGD as the optimizer, which converges slowly and is prone to getting stuck in local minima.
2. **3 Layers, Adam Optimizer**: Switching to the Adam optimizer significantly improved the accuracy to **72%**, as Adam adapts the learning rate for each parameter, resulting in faster convergence.

3. **Dropout Layer Addition**: Adding a **Dropout layer (0.3)** to combat overfitting further increased testing accuracy to **75%**, suggesting that regularization through dropout helped in preventing overfitting.
4. **L2 Regularization**: The combination of **L2 regularization** and dropout pushed the accuracy to **78%**. L2 regularization penalizes large weights, making the model less prone to overfitting.
5. **L1 Regularization + Batch Normalization**: L1 regularization combined with **batch normalization** led to stable training and a **76%** testing accuracy. Batch normalization helped the network stabilize the training process.
6. **Adam with Batch Normalization**: Using Adam with **batch normalization** and increasing the epochs to 100 improved accuracy to **81%**. Batch normalization reduces internal covariate shift, helping the model converge better with Adam's adaptive learning.
7. **4 Layers + Dropout + L2**: Finally, increasing the complexity of the architecture to 4 layers, with **dropout** and **L2 regularization**, led to the best result of **85% accuracy**. This configuration balanced regularization with model complexity, preventing overfitting while leveraging the additional layers.

5. Conclusion

The results show that increasing the model complexity, applying regularization techniques like dropout and L2, and using the Adam optimizer with batch normalization yielded the best performance. The combination of **4 layers**, **dropout**, and **L2 regularization** was particularly effective, achieving the highest accuracy of **85%**.
