Regularized Neural Network Model

Kazeem Adesina DAUDA

AI4WIA BOOTCAMP GROUP

September 8, 2023

Introduction

- Neural networks are powerful machine learning models.
- Regularization techniques help prevent overfitting.
- In this presentation, we'll focus on L2 regularization.

Neural Network Architecture

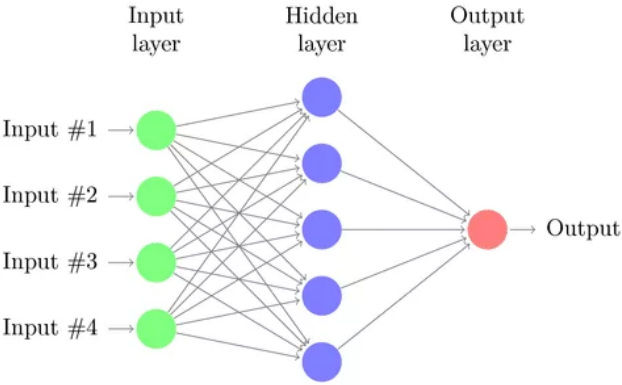

Figure: A simple neural network architecture.

Mathematical Formulation

Forward Pass:

1. Calculate the output of the hidden layer:

$$Z_1 = X \cdot W_1 + b_1$$
$$A_1 = \text{activation}(Z_1)$$

2. Calculate the output of the output layer (predictions):

$$Z_2 = A_1 \cdot W_2 + b_2$$
$$Y = \text{activation}(Z_2)$$

Loss Function with L2 Regularization

For Regression (MSE with L2 Regularization):

$$L = \frac{1}{N} \sum_{i=1}^{N} \left(\frac{1}{2} ||Y_i - Y_{\text{true}_i}||^2 + \frac{\lambda}{2} (||W_1||^2 + ||W_2||^2) \right)$$

For Classification (Cross-Entropy with L2 Regularization):

$$L = -\frac{1}{N} \sum_{i=1}^{N} (Y_{\text{true}_i} \cdot \log(Y_i) + (1 - Y_{\text{true}_i}) \cdot \log(1 - Y_i)) + \frac{\lambda}{2} (||W_1||^2 + ||W_2||^2$$

Where:

- N is the number of training examples.
- λ is the regularization parameter.
- `activation` is the chosen activation function (e.g., ReLU, sigmoid, softmax).

Training Process

- Minimize the regularized loss function using optimization algorithms (e.g., gradient descent).
- Update weights and biases to reduce the loss.
- Regularization terms penalize large weights.

Conclusion

- Regularization is crucial for preventing overfitting in neural networks.
- L2 regularization (weight decay) is a common technique.
- Fine-tune the regularization parameter λ for best results.

Questions?

Thank you for your attention. Any questions?