- **Input Email Example:**
  - **X1 = 4** (4 advertisement-related words)
  - **X2 = 3** (3 exclamation marks)
  - **X3 = 1** (1 suspicious link)
- **Network Structure:**
  - **Input Layer:** 3 neurons (X1, X2, X3)
  - **Hidden Layer:** 2 neurons
  - **Output Layer:** 1 neuron (binary classification: spam or not spam)
- **Weights and Biases:**
  - **Hidden Layer:**
    - $W_{11} = 0.2, W_{12} = -0.3, W_{13} = 0.4$
    - $W_{21} = 0.1, W_{22} = 0.6, W_{23} = -0.4$
    - $b_1^{(1)} = 0.1, b_2^{(1)} = -0.2$
  - **Output Layer:**
    - $W_1 = 0.3, W_2 = -0.5$
    - $b^{(2)} = 0.2$

Forward Propagation

1. Hidden Layer Calculations:

For **Hidden Neuron 1**:

$$Z_1^{(1)} = (0.2 \times 4) + (-0.3 \times 3) + (0.4 \times 1) + 0.1$$
$$Z_1^{(1)} = 0.8 - 0.9 + 0.4 + 0.1 = 0.4$$

Apply the ReLU activation function:

$$A_1^{(1)} = \text{ReLU}(Z_1^{(1)}) = \text{ReLU}(0.4) = 0.4$$

For **Hidden Neuron 2**:

$$Z_2^{(1)} = (0.1 \times 4) + (0.6 \times 3) + (-0.4 \times 1) - 0.2$$
$$Z_2^{(1)} = 0.4 + 1.8 - 0.4 - 0.2 = 1.6$$

Apply the ReLU activation function:

$$A_2^{(1)} = \text{ReLU}(Z_2^{(1)}) = \text{ReLU}(1.6) = 1.6$$

2. Output Layer Calculation:

$$Z^{(2)} = (0.3 \times 0.4) + (-0.5 \times 1.6) + 0.2$$
$$Z^{(2)} = 0.12 - 0.8 + 0.2 = -0.48$$

Apply the Sigmoid activation function:

$$\hat{y} = \sigma(Z^{(2)}) = \frac{1}{1 + e^{-(-0.48)}} \approx 0.382$$

Loss Calculation

Assume the true label $y = 1$ (indicating that the email is spam).

Using Binary Cross-Entropy Loss:

$$\text{Loss} = -[y \cdot \log(\hat{y}) + (1 - y) \cdot \log(1 - \hat{y})]$$
$$\text{Loss} = -[1 \cdot \log(0.382) + 0 \cdot \log(1 - 0.382)]$$

$$\text{Loss} = -\log(0.382) \approx 0.961$$

Backpropagation

We'll calculate the gradients for one iteration and update the weights accordingly.

1. Compute the Gradient of the Loss with respect to $Z^{(2)}$:

$$\frac{\partial \text{Loss}}{\partial Z^{(2)}} = \hat{y} - y = 0.382 - 1 = -0.618$$

2. Compute the Gradient with respect to Output Layer Weights:

For W_1:

$$\frac{\partial \text{Loss}}{\partial W_1} = \frac{\partial \text{Loss}}{\partial Z^{(2)}} \cdot \frac{\partial Z^{(2)}}{\partial W_1} = -0.618 \times A_1^{(1)} = -0.618 \times 0.4 = -0.2472$$

For W_2:

$$\frac{\partial \text{Loss}}{\partial W_2} = \frac{\partial \text{Loss}}{\partial Z^{(2)}} \cdot \frac{\partial Z^{(2)}}{\partial W_2} = -0.618 \times A_2^{(1)} = -0.618 \times 1.6 = -0.9888$$

For $b^{(2)}$ (bias in the output layer):

$$\frac{\partial \text{Loss}}{\partial b^{(2)}} = \frac{\partial \text{Loss}}{\partial Z^{(2)}} \times 1 = -0.618$$

3. Update Weights and Biases (using a learning rate $\alpha = 0.01$):

For W_1:

$$W_1^{\text{new}} = W_1^{\text{old}} - \alpha \times \frac{\partial \text{Loss}}{\partial W_1} = 0.3 - 0.01 \times (-0.2472) = 0.3 + 0.002472 = 0.302472$$

For W_2:

$$W_2^{\text{new}} = W_2^{\text{old}} - \alpha \times \frac{\partial \text{Loss}}{\partial W_2} = -0.5 - 0.01 \times (-0.9888) = -0.5 + 0.009888 = -0.490112$$

For $b^{(2)}$:

$$b^{\text{new}} = b^{\text{old}} - \alpha \times \frac{\partial \text{Loss}}{\partial b^{(2)}} = 0.2 - 0.01 \times (-0.618) = 0.2 + 0.00618 = 0.20618$$

4. Gradients for the Hidden Layer:

First, we calculate the gradients for $Z_1^{(1)}$ and $Z_2^{(1)}$.

For $Z_1^{(1)}$:

$$\frac{\partial \text{Loss}}{\partial Z_1^{(1)}} = \frac{\partial \text{Loss}}{\partial Z^{(2)}} \times W_1 \times \text{ReLU}'(Z_1^{(1)})$$

Since $\text{ReLU}'(Z_1^{(1)}) = 1$ (as $Z_1^{(1)} > 0$),

$$\frac{\partial \text{Loss}}{\partial Z_1^{(1)}} = -0.618 \times 0.3 = -0.1854$$

For $Z_2^{(1)}$:

$$\frac{\partial \text{Loss}}{\partial Z_2^{(1)}} = \frac{\partial \text{Loss}}{\partial Z^{(2)}} \times W_2 \times \text{ReLU}'(Z_2^{(1)})$$

$$\frac{\partial \text{Loss}}{\partial Z_2^{(1)}} = -0.618 \times (-0.5) = 0.309$$

5. Gradient with respect to Weights W_{11}, W_{12}, W_{13}:

For W_{11}:

$$\frac{\partial \text{Loss}}{\partial W_{11}} = \frac{\partial \text{Loss}}{\partial Z_1^{(1)}} \cdot \frac{\partial Z_1^{(1)}}{\partial W_{11}} = -0.1854 \times X_1 = -0.1854 \times 4 = -0.7416$$

For W_{12}:

$$\frac{\partial \text{Loss}}{\partial W_{12}} = -0.1854 \times X_2 = -0.1854 \times 3 = -0.5562$$

For W_{13}:

$$\frac{\partial \text{Loss}}{\partial W_{13}} = -0.1854 \times X_3 = -0.1854 \times 1 = -0.1854$$

6. Update the Weights W_{11}, W_{12}, W_{13} and Bias $b_1^{(1)}$:

For W_{11}:

$$W_{11}^{\text{new}} = W_{11}^{\text{old}} - \alpha \times \frac{\partial \text{Loss}}{\partial W_{11}} = 0.2 - 0.01 \times (-0.7416) = 0.2 + 0.007416 = 0.207416$$

For W_{12}:

$$W_{12}^{\text{new}} = W_{12}^{\text{old}} - \alpha \times \frac{\partial \text{Loss}}{\partial W_{12}} = -0.3 - 0.01 \times (-0.5562) = -0.3 + 0.005562 = -0.294438$$

For W_{13}:

$$W_{13}^{\text{new}} = W_{13}^{\text{old}} - \alpha \times \frac{\partial \text{Loss}}{\partial W_{13}} = 0.4 - 0.01 \times (-0.1854) = 0.4 + 0.001854 = 0.401854$$

For $b_1^{(1)}$:

$$b_1^{\text{new}} = b_1^{\text{old}} - \alpha \times \frac{\partial \text{Loss}}{\partial b_1^{(1)}} = 0.1 - 0.01 \times (-0.1854) = 0.1 + 0.001854 = 0.101854$$

[redacted]

Backpropagation for the Second Hidden Neuron

1. Gradient with respect to Weights W_{21}, W_{22}, W_{23}:

We already calculated the gradient for $Z_2^{(1)}$ as:

$$\frac{\partial \text{Loss}}{\partial Z_2^{(1)}} = 0.309$$

Now, let's calculate the gradients for the weights associated with the second hidden neuron.

For W_{21}:

$$\frac{\partial \text{Loss}}{\partial W_{21}} = \frac{\partial \text{Loss}}{\partial Z_2^{(1)}} \cdot \frac{\partial Z_2^{(1)}}{\partial W_{21}} = 0.309 \times X_1 = 0.309 \times 4 = 1.236$$

For W_{22}:

$$\frac{\partial \text{Loss}}{\partial W_{22}} = \frac{\partial \text{Loss}}{\partial Z_2^{(1)}} \cdot \frac{\partial Z_2^{(1)}}{\partial W_{22}} = 0.309 \times X_2 = 0.309 \times 3 = 0.927$$

For W_{23}:

$$\frac{\partial \text{Loss}}{\partial W_{23}} = \frac{\partial \text{Loss}}{\partial Z_2^{(1)}} \cdot \frac{\partial Z_2^{(1)}}{\partial W_{23}} = 0.309 \times X_3 = 0.309 \times 1 = 0.309$$

2. Update the Weights W_{21}, W_{22}, W_{23} and Bias $b_2^{(1)}$:

Using the learning rate $\alpha = 0.01$:

For W_{21}:

$$W_{21}^{\text{new}} = W_{21}^{\text{old}} - \alpha \times \frac{\partial \text{Loss}}{\partial W_{21}} = 0.1 - 0.01 \times 1.236 = 0.1 - 0.01236 = 0.08764$$

For W_{22}:

$$W_{22}^{\text{new}} = W_{22}^{\text{old}} - \alpha \times \frac{\partial \text{Loss}}{\partial W_{22}} = 0.6 - 0.01 \times 0.927 = 0.6 - 0.00927 = 0.59073$$

For W_{23}:

$$W_{23}^{\text{new}} = W_{23}^{\text{old}} - \alpha \times \frac{\partial \text{Loss}}{\partial W_{23}} = -0.4 - 0.01 \times 0.309 = -0.4 - 0.00309 = -0.40309$$

For $b_2^{(1)}$:

$$b_2^{\text{new}} = b_2^{\text{old}} - \alpha \times \frac{\partial \text{Loss}}{\partial b_2^{(1)}} = -0.2 - 0.01 \times 0.309 = -0.2 - 0.00309 = -0.20309$$

Summary of Updated Weights and Biases

After one iteration, the updated weights and biases are as follows:

- **Hidden Layer 1:**
  - W_{11}: 0.207416
  - W_{12}: -0.294438
  - W_{13}: 0.401854
  - $b_1^{(1)}$: 0.101854
- **Hidden Layer 2:**
  - W_{21}: 0.08764
  - W_{22}: 0.59073
  - W_{23}: -0.40309
  - $b_2^{(1)}$: -0.20309
- **Output Layer:**
  - W_1: 0.302472
  - W_2: -0.490112
  - $b^{(2)}$: 0.20618