```
import numpy as np
import matplotlib.pyplot as plt
# Sigmoid activation and derivative
def sigmoid(x):
    return 1 / (1 + np.exp(-x))
def sigmoid_derivative(x):
   return x * (1 - x) # derivative assumes input = sigmoid(x)
# Training data (XOR problem)
X = np.array([[0,0],[0,1],[1,0],[1,1]])
y = np.array([[0],[1],[1],[0]])
# Set random seed for reproducibility
np.random.seed(42)
# Network architecture
input size = 2
hidden_size = 2
output_size = 1
# Initialize weights and biases
W1 = np.random.randn(input_size, hidden_size)
b1 = np.zeros((1, hidden_size))
W2 = np.random.randn(hidden size, output size)
b2 = np.zeros((1, output_size))
# Hyperparameters
1r = 0.1
epochs = 10000
losses = []
# Training loop
for epoch in range(epochs):
   # ---- Forward pass ----
    z1 = np.dot(X, W1) + b1
    a1 = sigmoid(z1)
    z2 = np.dot(a1, W2) + b2
    a2 = sigmoid(z2)
    # ---- Loss (Mean Squared Error) ----
    loss = np.mean((y - a2) ** 2)
    losses.append(loss)
    # ---- Backpropagation ----
   d_a2 = (a2 - y)
d_z2 = d_a2 * sigmoid_derivative(a2)
    dW2 = np.dot(a1.T, d_z2)
    db2 = np.sum(d_z2, axis=0, keepdims=True)
    d_a1 = np.dot(d_z2, W2.T)
    d_z1 = d_a1 * sigmoid_derivative(a1)
    dW1 = np.dot(X.T, d_z1)
    db1 = np.sum(d_z1, axis=0, keepdims=True)
    # ---- Update weights ----
    W2 -= 1r * dW2
    b2 -= 1r * db2
    W1 -= lr * dW1
    b1 -= lr * db1
    # Print loss occasionally
    if epoch % 2000 == 0:
        print(f"Epoch {epoch}, Loss: {loss:.4f}")
# ---- Final predictions ---
print("\nFinal Predictions:")
print(a2)
# ---- Plot the loss curve ----
plt.figure(figsize=(8,5))
plt.plot(losses, label="Training Loss", color="blue")
plt.xlabel("Epochs")
plt.ylabel("Loss")
plt.title("Loss Curve - Gradient Descent & Backpropagation")
plt.legend()
plt.grid(True)
plt.show()
```



