Customer Churn Prediction (Neural Network From Scratch)

By Muddasar Razzaq

Project Overview

- - Predict customer churn (Exit or Stay)
- - Manual preprocessing, encoding, scaling
- - Forward pass, backpropagation, gradient descent
- - Evaluate on test dataset

Dataset Description

- - 10,000 customer records
- - Features: CreditScore, Geography, Gender, Age,
- Tenure, Balance, Products, Credit Card, Activity, Salary
- - Target: Exited (1 = churn, 0 = stay)
- - Imbalanced (approx. 20% churn)

Data Preprocessing

- - Drop ID-like columns
- - Encode Gender manually (0/1)
- - One-hot encode Geography
- - Standardize numerical features
- - Manual train-test split (80/20)

Neural Network Architecture

- - Architecture: Input → 16 → 8 → Output
- - Hidden Layers: ReLU activation
- - Output Layer: Softmax
- - Loss: Categorical Cross-Entropy
- - Optimization: Gradient Descent

Forward & Backpropagation

- Forward Pass:
- - Z1 = ReLU(XW1 + b1)
- - Z2 = Softmax(Z1W2 + b2)

- Backpropagation:
- - Compute gradients for W and b
- - Update parameters using Gradient Descent

Training Loss Curve Insights

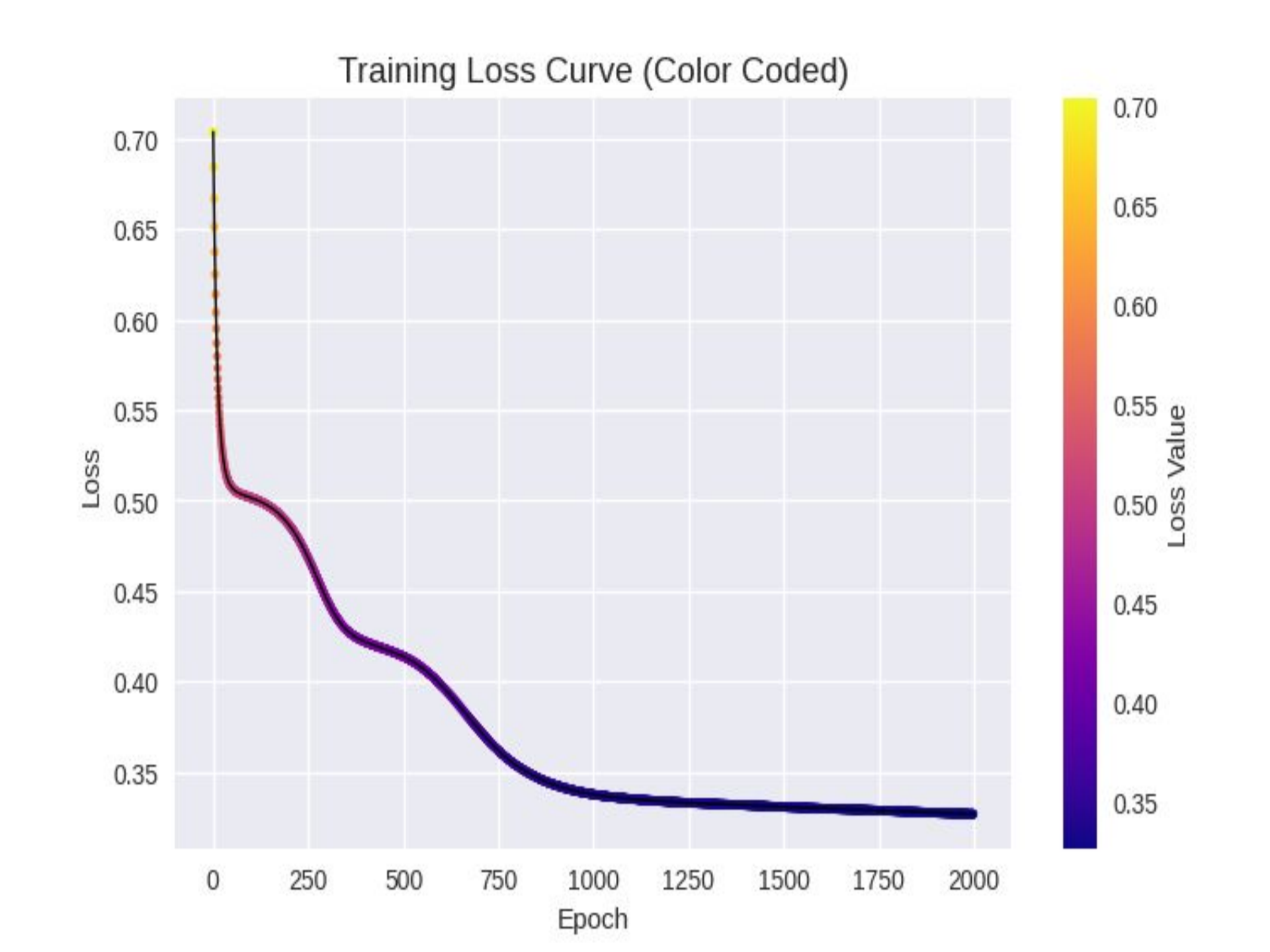
- - Loss decreases from ~0.70 → ~0.33
- - Smooth convergence
- - No overfitting observed
- - Indicates stable learning

```
Epoch 200/2000 - Loss: 0.4867
Epoch 400/2000 - Loss: 0.4220
Epoch 600/2000 - Loss: 0.3980
Epoch 800/2000 - Loss: 0.3533
Epoch 1000/2000 - Loss: 0.3376
Epoch 1200/2000 - Loss: 0.3335
Epoch 1400/2000 - Loss: 0.3316
Epoch 1600/2000 - Loss: 0.3298
Epoch 1800/2000 - Loss: 0.3283
Epoch 2000/2000 - Loss: 0.3267
```

Model Performance

- Training Accuracy: 86.3%
- Test Accuracy: 85.7%

- Interpretation:
- - Good generalization
- - Solid baseline for a manually built ANN


Training Loss Curve (Color Coded)
Loss
Epoch
Loss Value
0.70
0.65
0.60
0.55
0.50
0.45
0.40
0.35
0
250
500
750
1000
1250
1500
1750
2000

Business Insights

- - Older customers have higher churn probability
- - German customers churn more
- - Active members churn less
- - High balance customers usually stay

Conclusion

- - Complete ML pipeline built manually
- - Strong performance and clear insights
- - Demonstrates deep understanding of ANN fundamentals
- - Ready to extend with tuning or deployment