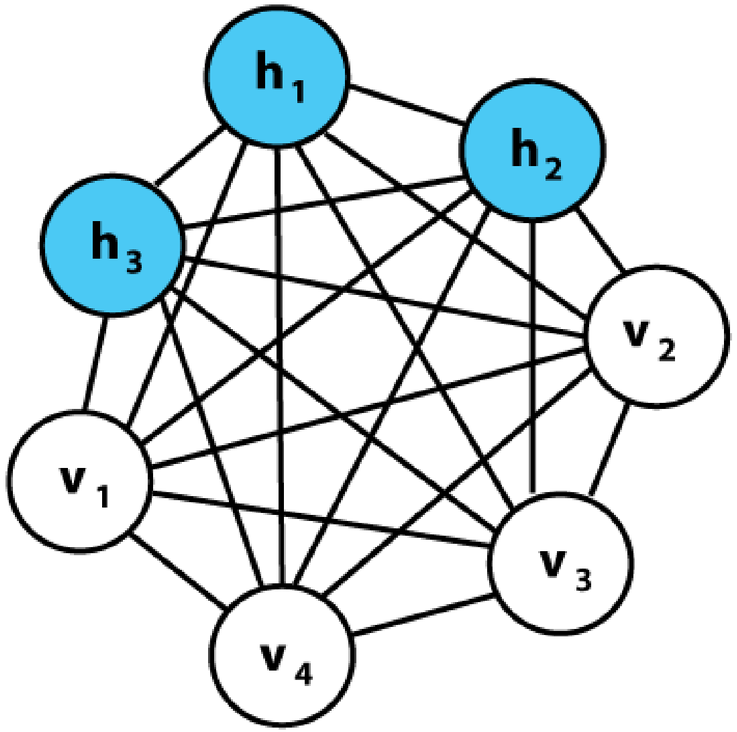
# Digit Classification using RBM as feature extractor

# demo

# Introduction to Boltzmann Machine

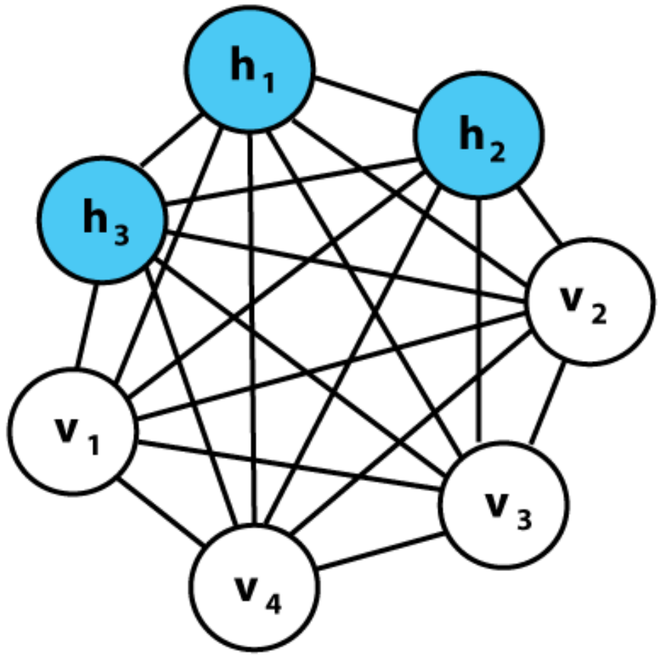
- A type of stochastic neural network used for probabilistic learning.
- Inspired by statistical mechanics (named after Ludwig Boltzmann).
- Learns complex data representations by modeling probability distributions.




$$E(v,h) = -\sum_i a_i v_i - \sum_j b_j h_j - \sum_i \sum_j v_i w_{i,j} h_j$$

# Architecture of Boltzmann Machine

- Fully connected network with:
  - Visible units (v) – Represent input data.
  - Hidden units (h) – Capture complex patterns.
- Each neuron is binary (0/1) and updates probabilistically.
- Bidirectional weighted connections (no layered structure like traditional NNs).

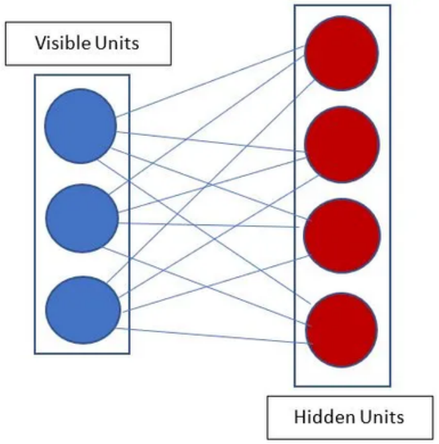


# Drawbacks of Boltzmann Machine

- Slow training: Requires Gibbs sampling, making large-scale learning inefficient.
- Computationally expensive: Partition function ZZZ is hard to compute.
- Fully connected structure: Leads to complex weight updates.
- Not practical for deep learning due to convergence issues.

# Introduction to Restricted Boltzmann Machine (RBM)

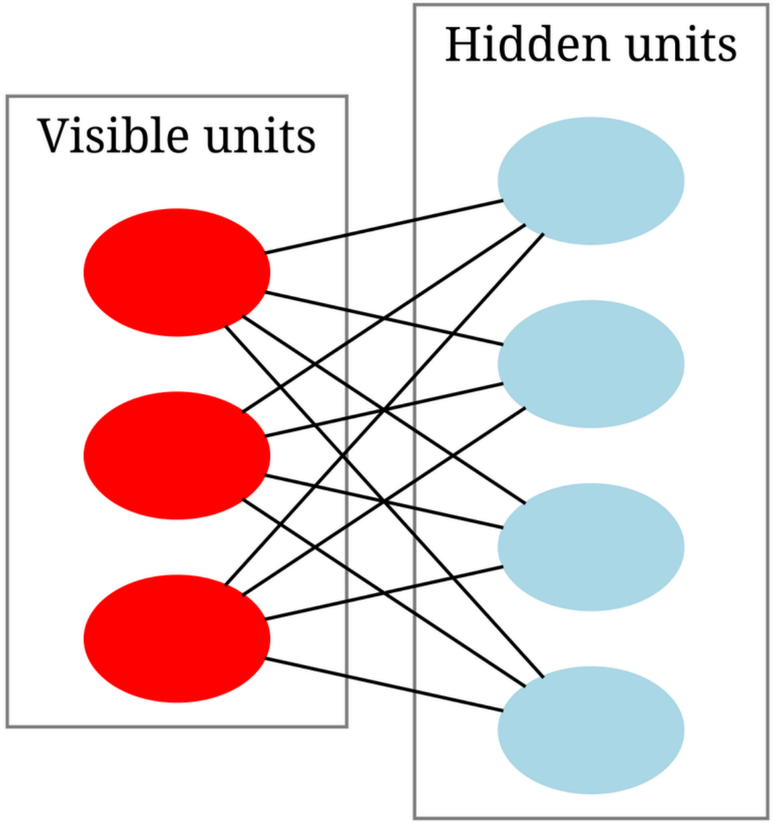
- A simplified version of Boltzmann Machine.
- Restricted architecture: No intra-layer connections → Faster training.
- Commonly used in dimensionality reduction, feature extraction, and recommendation systems.




$$E(v,h) = -\sum_{i \in visible} ai * vi - \sum_{j \in hidden} bj * hj - \sum_{i,j} vi * hj * wij$$

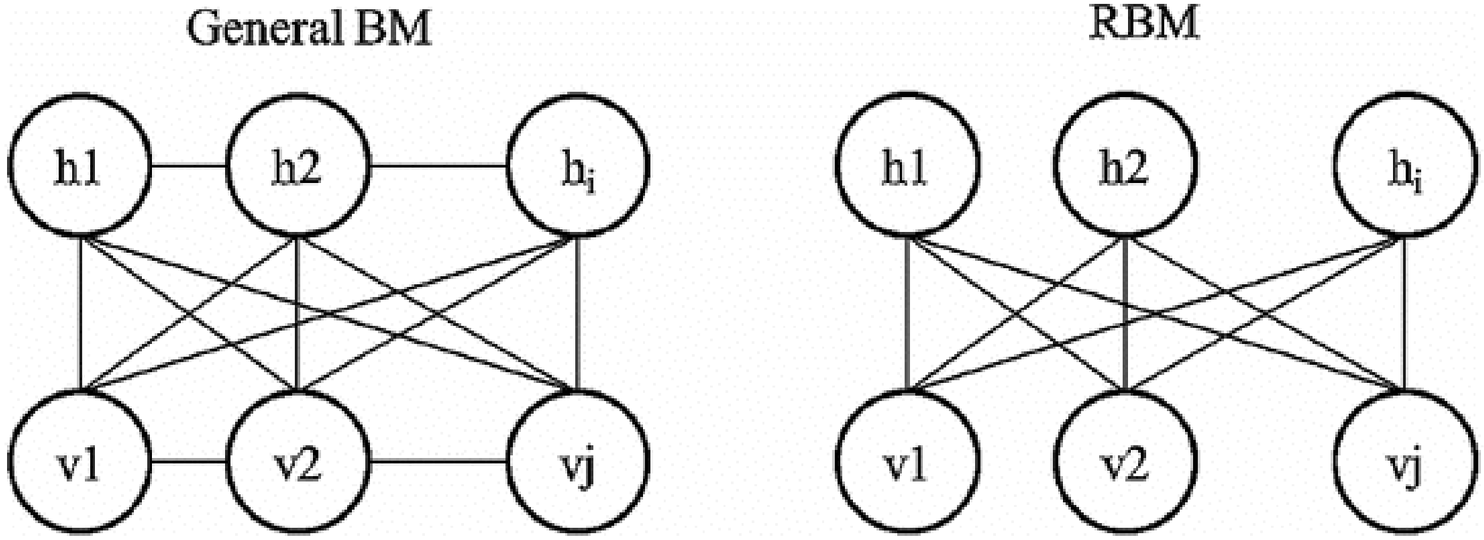
# Architecture of Restricted Boltzmann Machine (RBM)

- Two-layer structure:
  Visible layer (v) – Takes input data.
  Hidden layer (h) – Learns latent features.
- No connections within layers (unlike BM).
- Uses contrastive divergence (CD) for efficient training.



# Why RBM Over BM?

- Simpler Architecture: No intra-layer connections in RBM → Easier training.
- Faster Training: RBM uses Contrastive Divergence (CD), unlike BM's slow Gibbs sampling.
- Better Scalability: RBM is more efficient for large datasets and deep learning.
- Deep Learning Compatibility: RBM is used for pre-training deep networks; BM isn't practical.
- Real-World Applications: RBM is used in recommendation systems, feature extraction, and anomaly detection.



# Applications of RBM

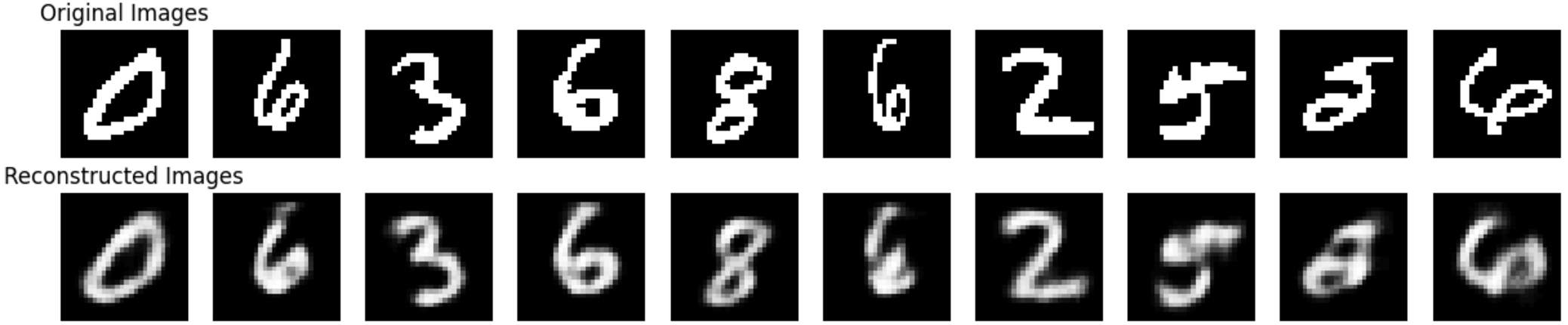
- Dimensionality reduction (like PCA).
- Feature learning (unsupervised pre-training in deep learning).
- Collaborative filtering (used in movie recommendation systems).
- Anomaly detection in cybersecurity.

# Our Project

# Methodology

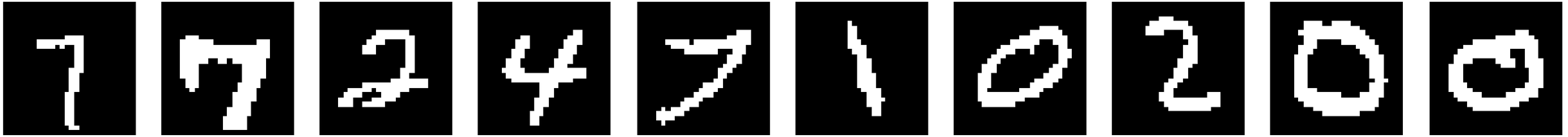
# Reconstruction After 0 Epochs


# Reconstruction After 2 Epochs

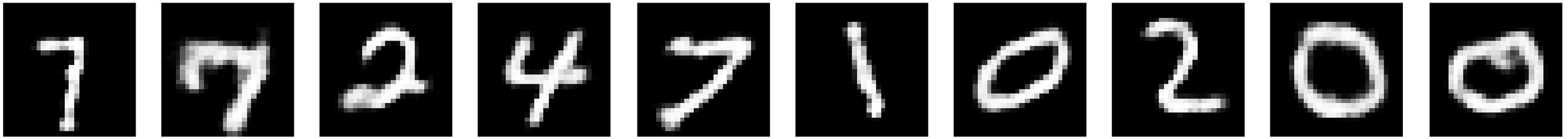


# Reconstruction After 10 Epochs

Original Images

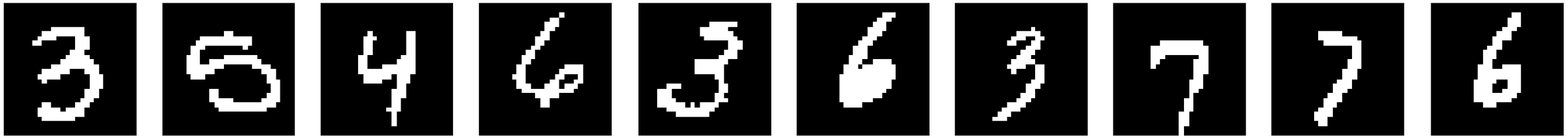


Reconstructed Images

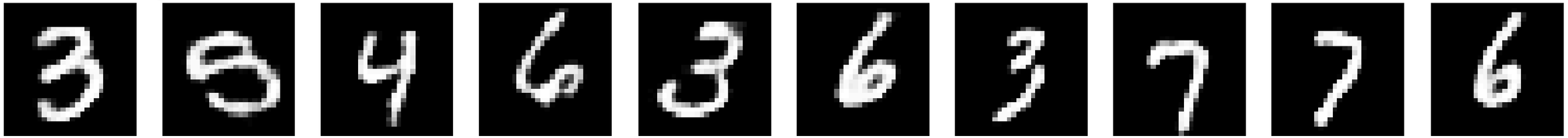


# Reconstruction After 20 Epochs

Original Images

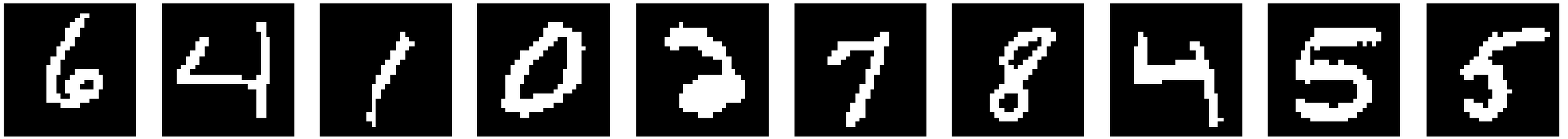


Reconstructed Images

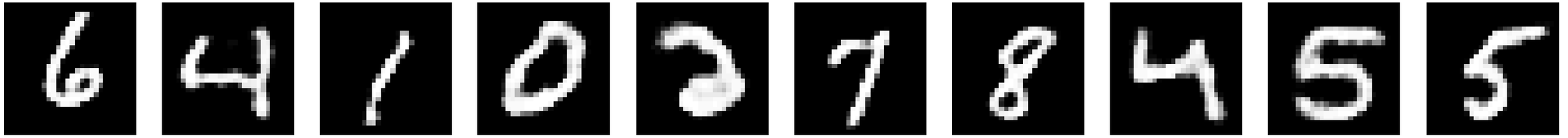


# Reconstruction After 30 Epochs

Original Images

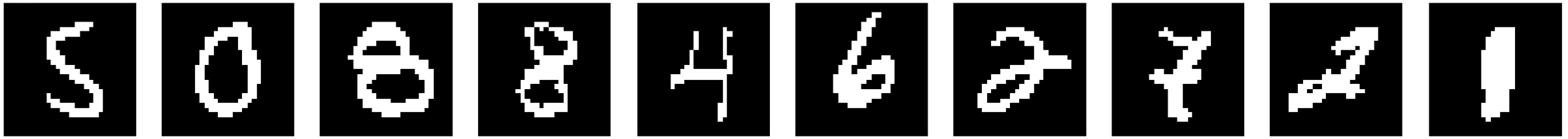


Reconstructed Images

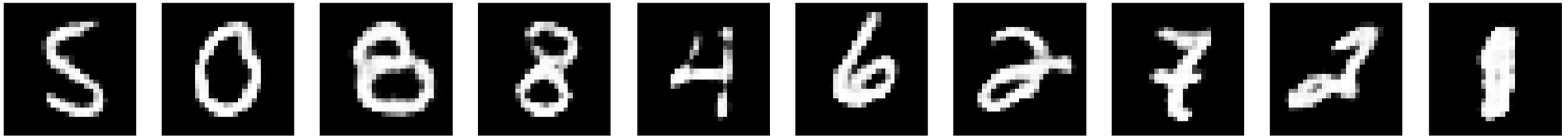


# Reconstruction After 50 Epochs

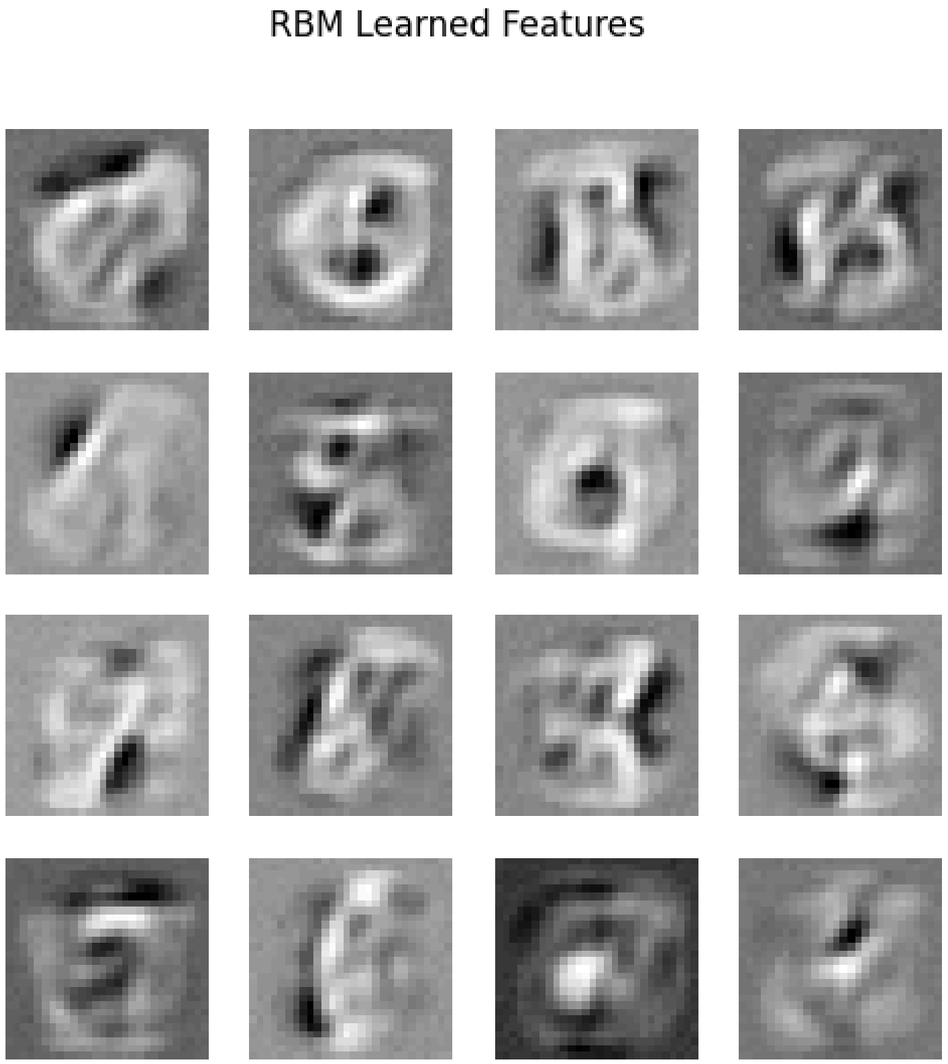
Original Images



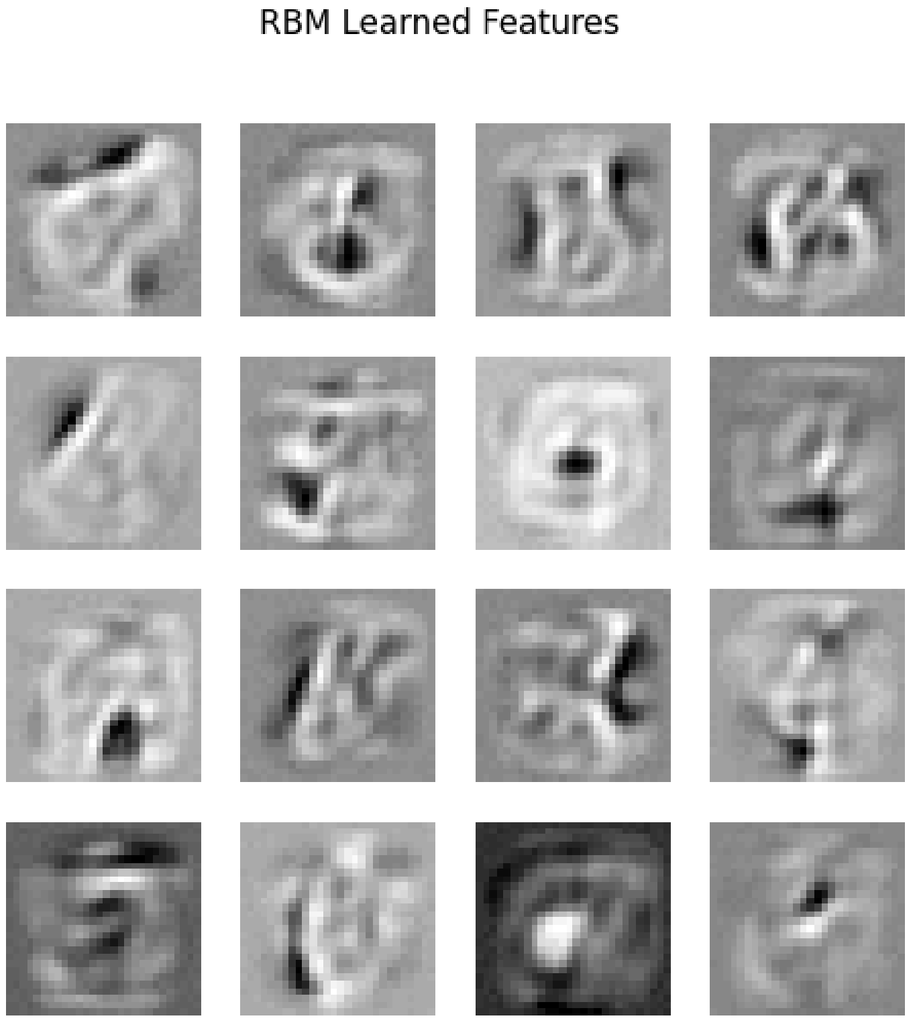
Reconstructed Images



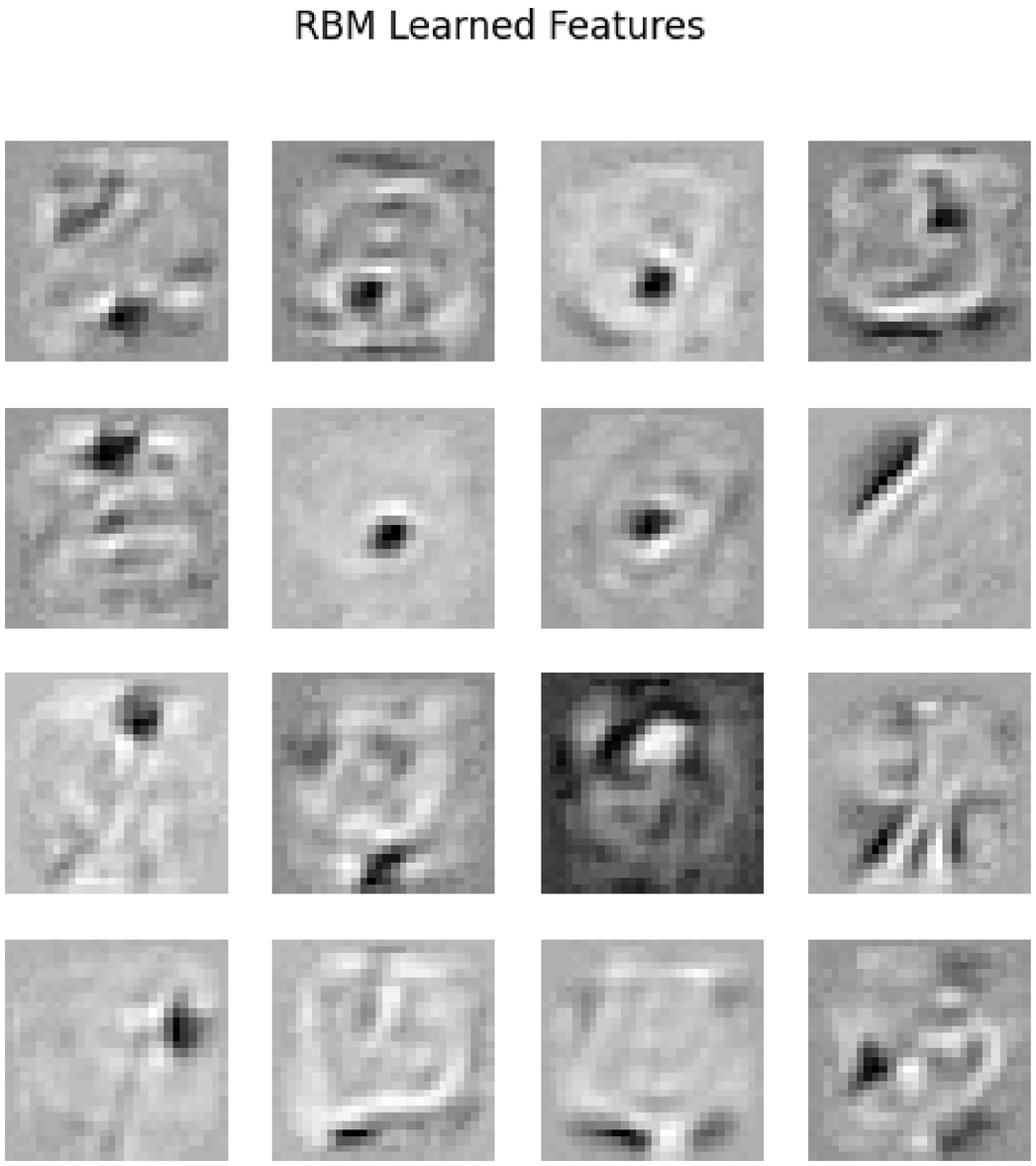
After 2 epochs



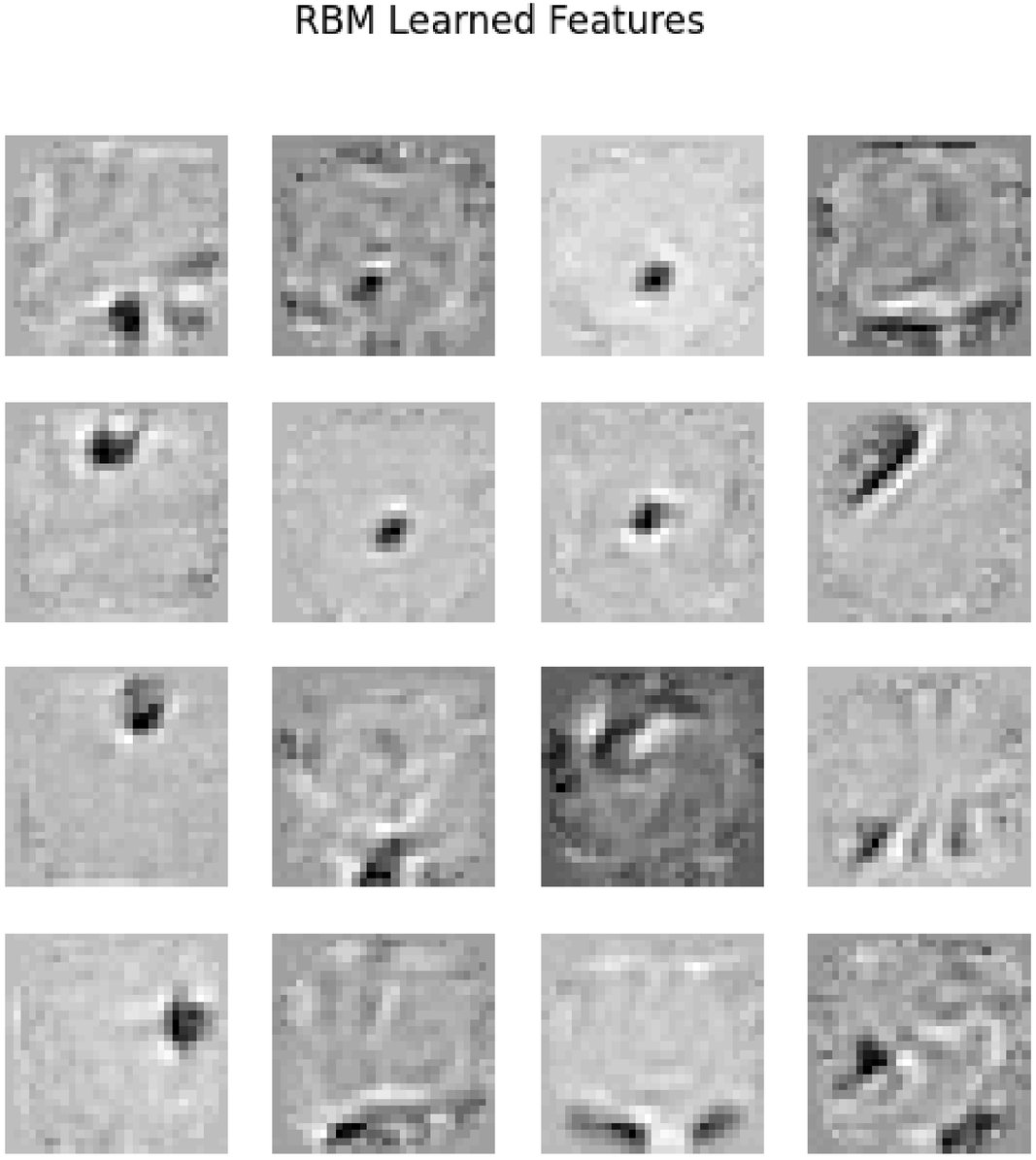

After 4 epochs



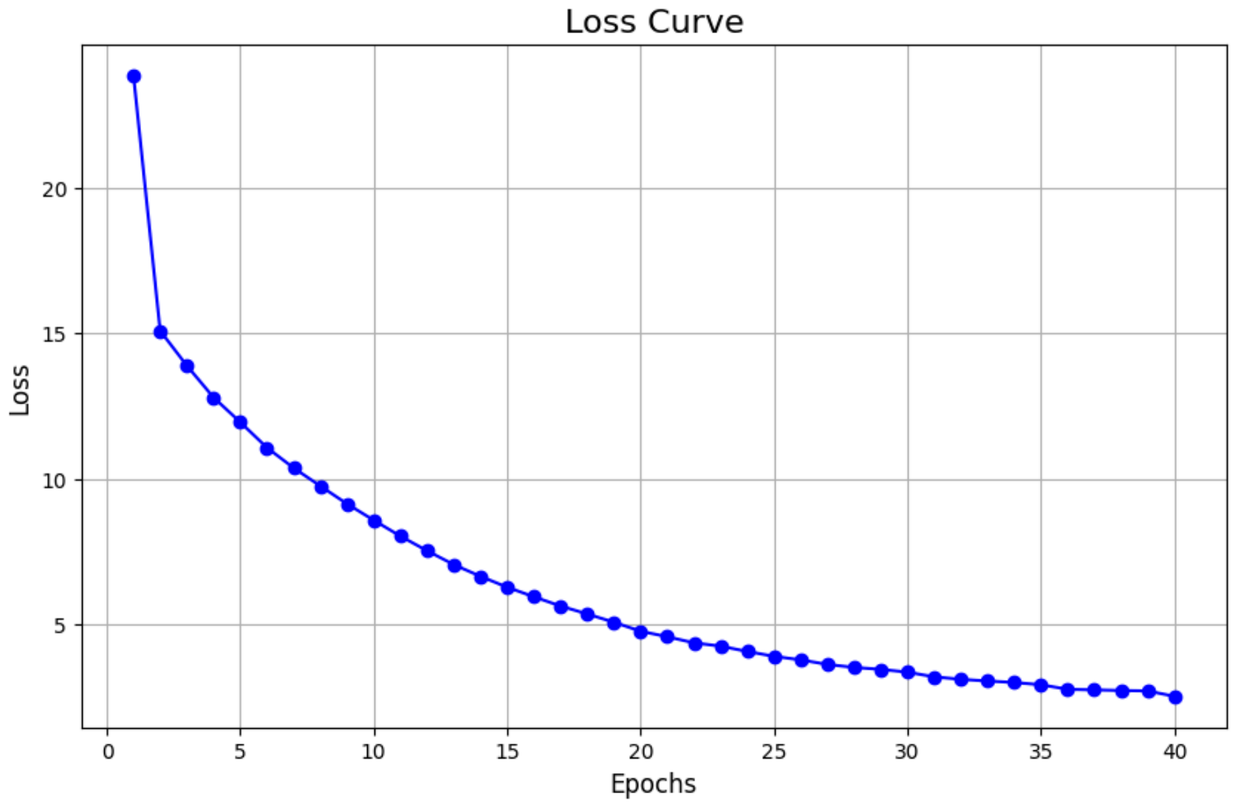
## After 10 epochs




## After 50 epochs



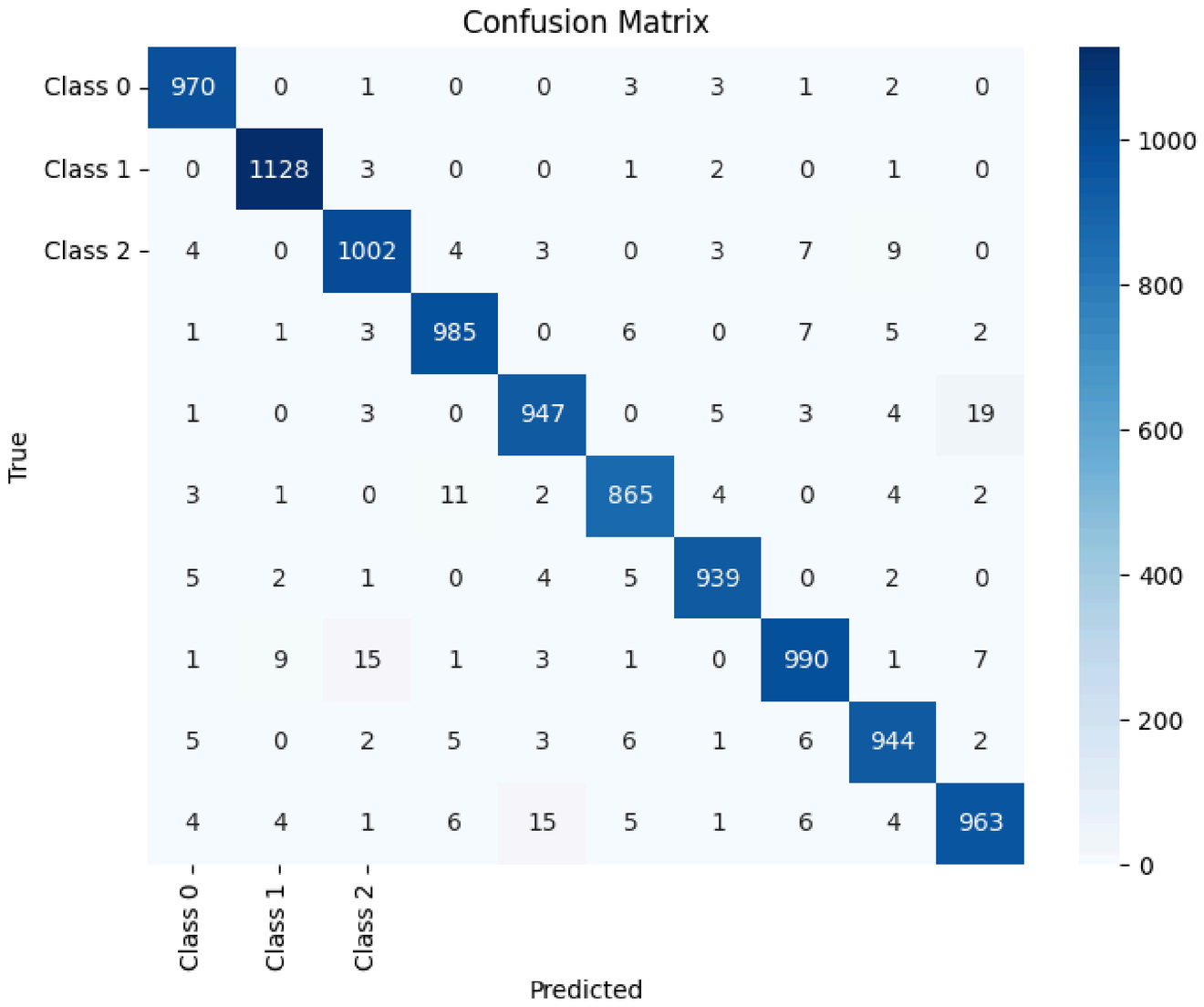
# Loss Curve



# Results

| Class | Precision | Recall | F1-Score | Support |
|---|---|---|---|---|
| 0 | 0.98 | 0.99 | 0.98 | 980 |
| 1 | 0.99 | 0.99 | 0.99 | 1135 |
| 2 | 0.97 | 0.97 | 0.97 | 1032 |
| 3 | 0.97 | 0.98 | 0.97 | 1010 |
| 4 | 0.97 | 0.96 | 0.97 | 982 |
| 5 | 0.97 | 0.97 | 0.97 | 892 |
| 6 | 0.98 | 0.98 | 0.98 | 958 |
| 7 | 0.97 | 0.96 | 0.97 | 1028 |
| 8 | 0.97 | 0.97 | 0.97 | 974 |
| 9 | 0.97 | 0.95 | 0.96 | 1009 |
| **Accuracy** | | | **0.97** | **10000** |
| **Macro avg** | 0.97 | 0.97 | 0.97 | 10000 |
| **Weighted avg** | 0.97 | 0.97 | 0.97 | 10000 |

# Confusion Matrix



# Thank You