Day 12, Training Neural Networks – Optimization Made Simple

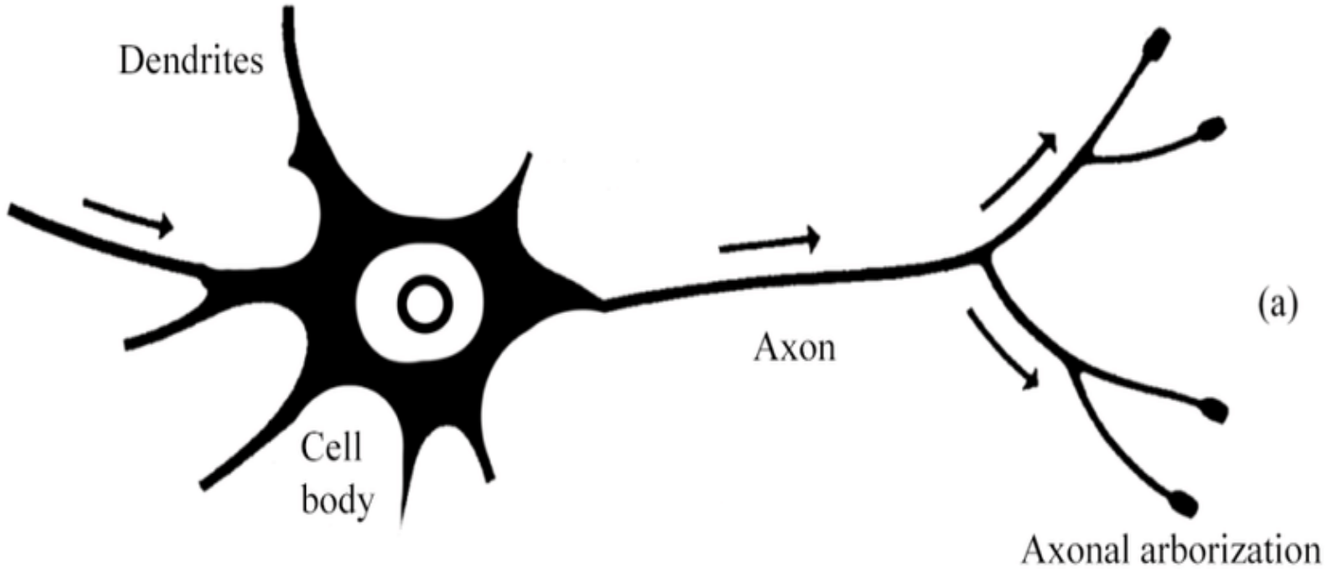
This presentation explores neural network training. We delve into optimization techniques. We simplify complex concepts. Our goal is to provide a clear understanding. We aim to make neural network training accessible.

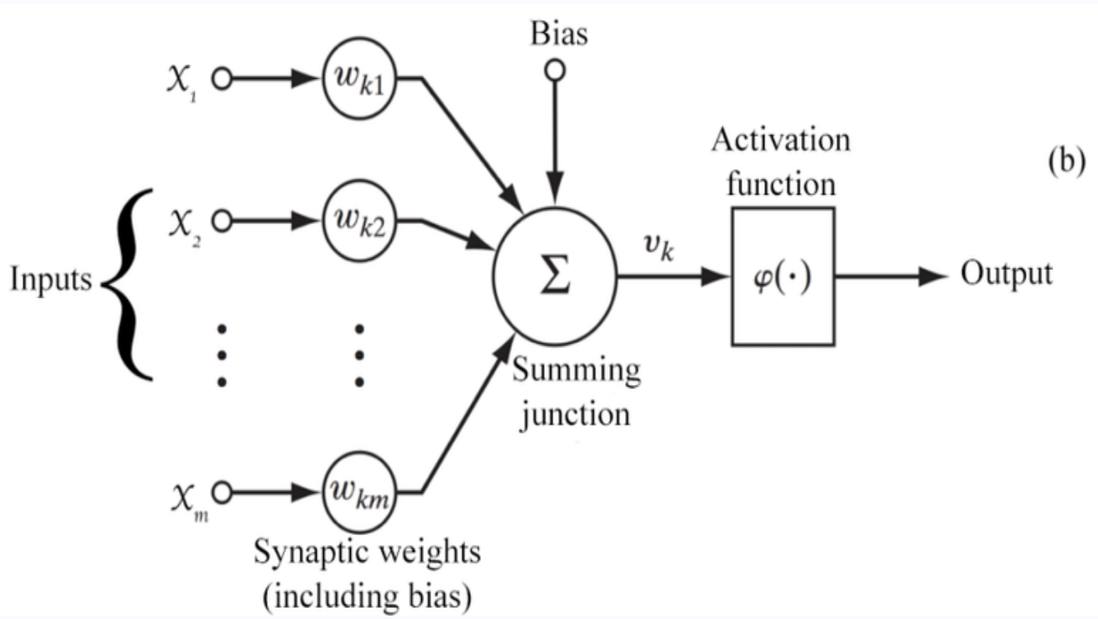
Presented By
Maheshkumar Paik

Agenda

- Introduction to Neural Networks
- The Training Process in Neural Networks
- Understanding Backpropagation
- Optimization Algorithms Overview
- Deep Dive: Gradient Descent
- Adam Optimizer: A Comprehensive Look
- How Adam Works: The Mechanics
- Benefits of Using Adam
- Conclusion: Mastering Optimization

Can Increase the speed to 1.5x





Introduction to Neural Networks

1 Mimicking the Human Brain

Neural networks are inspired by the human brain. They learn from data. They identify patterns.

2 Versatile Applications

They are used in image recognition. Also, they support speech-to-text. They perform predictive modeling tasks.

3 Key Components

Neurons (nodes), weights, biases, and layers are important. Input, hidden, and output layers form the structure.

The Training Process in Neural Networks

1

Forward Propagation

Data moves through the network.

2

Loss Calculation

Error between predictions and actuals.

3

Backpropagation

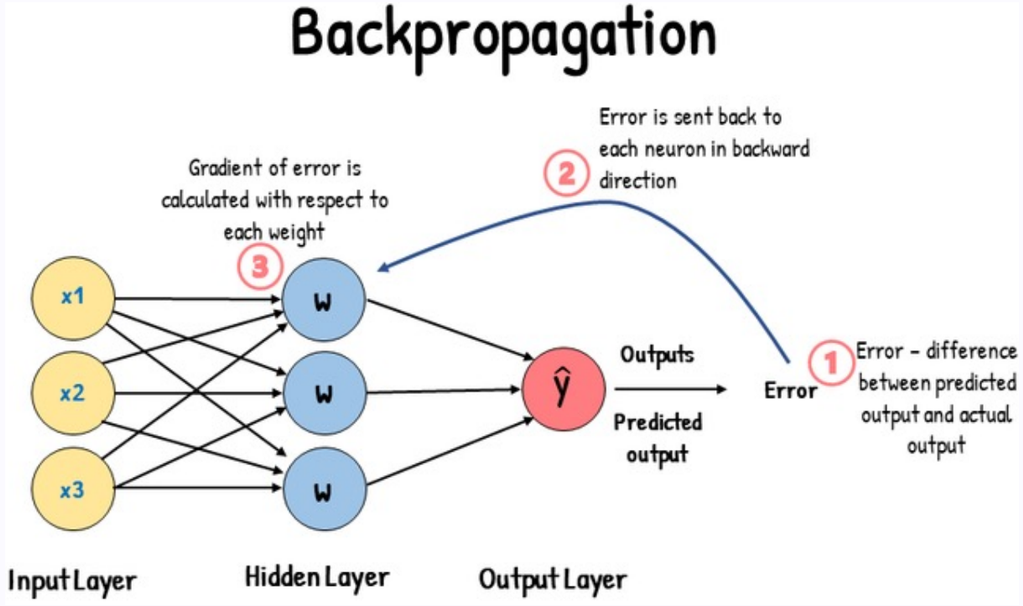
Adjusting weights based on error.

4

Optimization

Minimizing the error.

Backpropagation Explained



1

Compute Loss

Calculate the error.

2

Calculate Gradients

Determine error contribution.

3

Update Weights and Biases

Reduce error impact.

4

Repeat

Until convergence.

Introduction to Optimization Algorithms

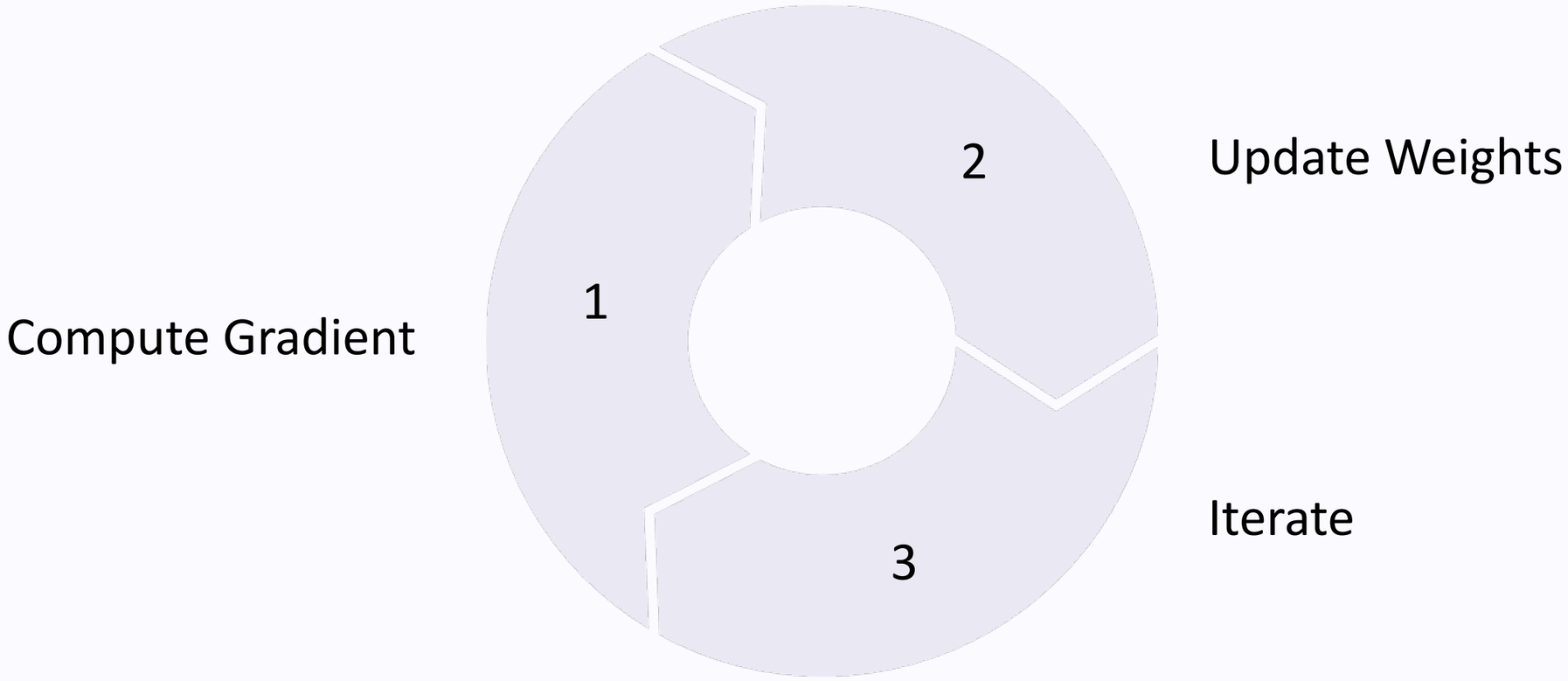
Gradient Descent

Basic, but slow optimization method.

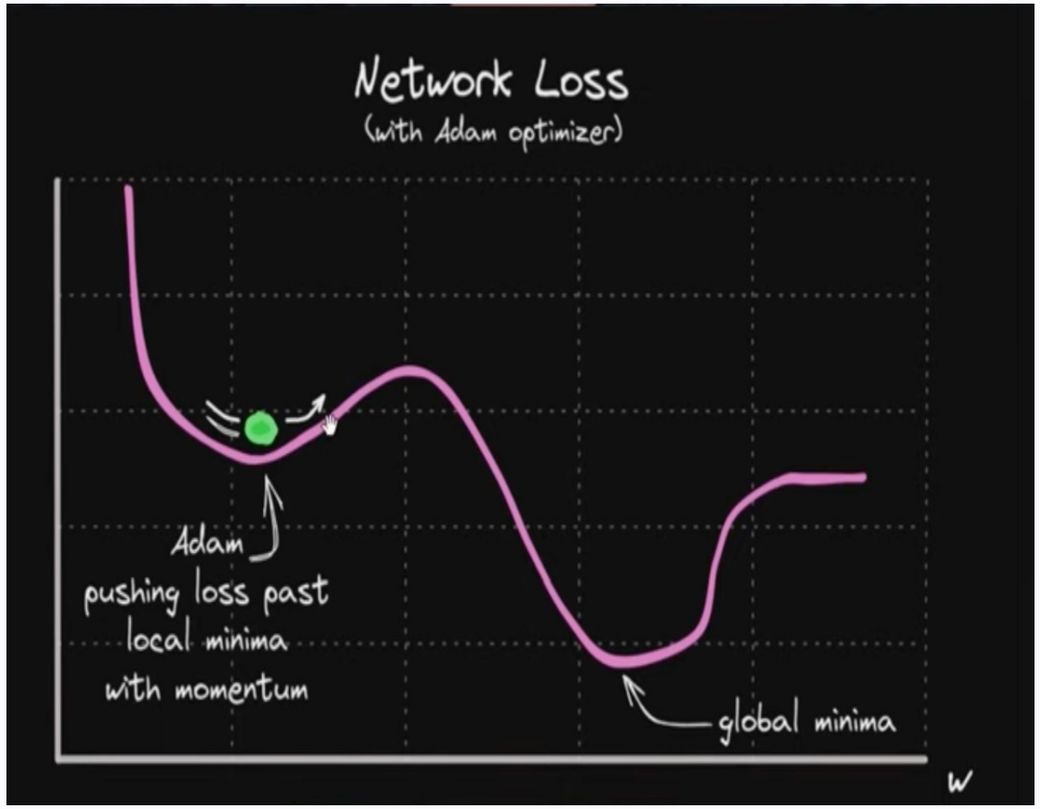
Adam

Combines best of both worlds. Widely used.

Gradient Descent: The Core Algorithm Explained



Gradient descent iteratively adjusts weights. It minimizes the loss function. The goal is to find the optimal parameters. This approach leads to better accuracy.



What is the Adam Optimizer?

Faster Learning

Faster, stable learning.

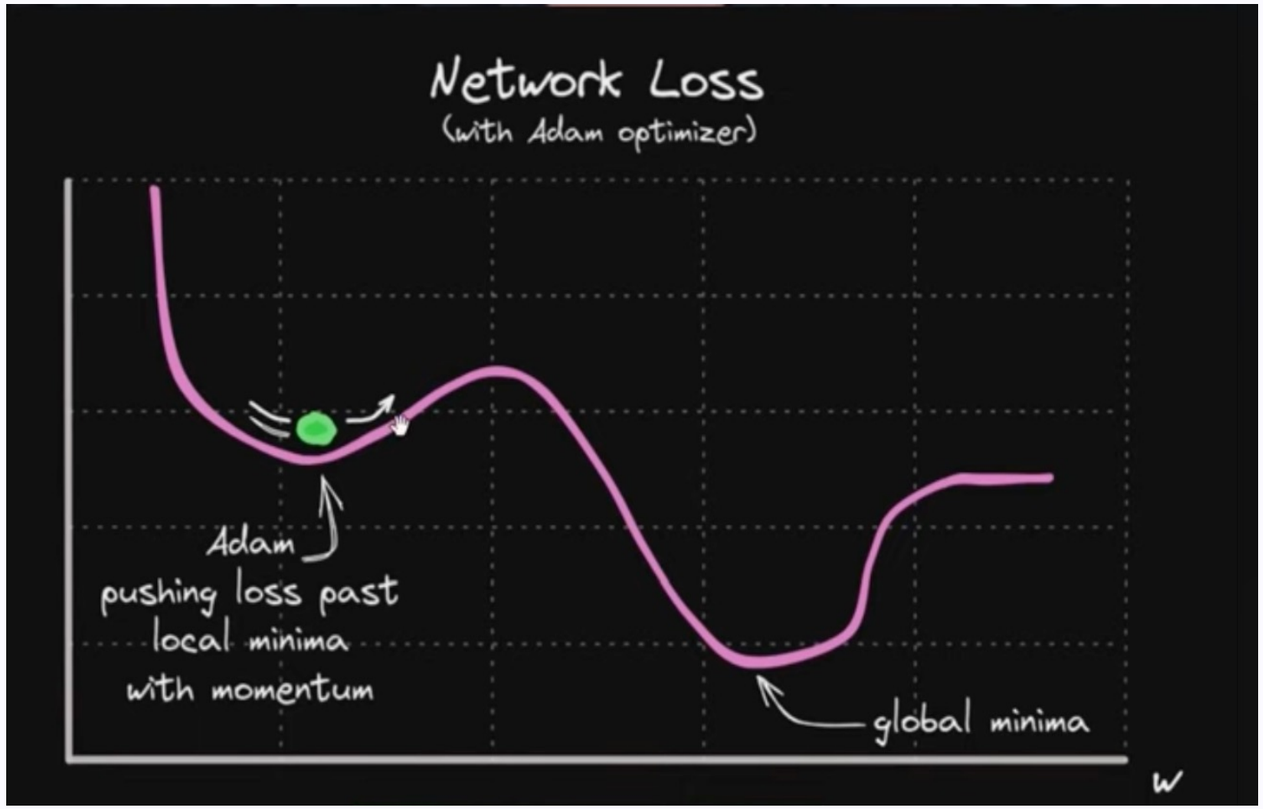
Adaptive Rate

Adaptive learning rate for each parameter.

Large Datasets

Efficient for large datasets.

How Adam Optimizer Works



Adam

Combines Momentum and RMSprop.

RMSprop

Scales learning rate.

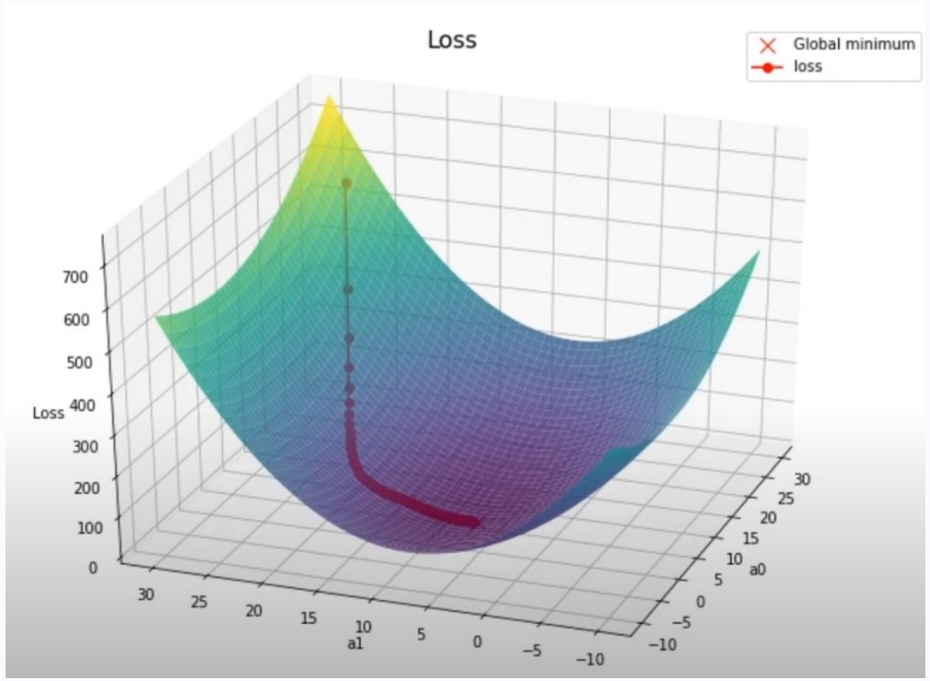
Momentum

Tracks past gradients.

The Adam optimizer is widely used. It keeps track of past gradients. It also scales learning rates. Adam maintains moving averages for optimization.

Benefits of Using Adam

Fast Convergence	Reduces training time effectively.
Stable Learning	Handles noisy data. Adapts to landscapes.
Widely Used	Works for most architectures.



Conclusion: Mastering Neural Network Optimization

Training involves Forward Propagation. This step is followed by Loss Calculation. Then Backpropagation. The process ends with Optimization using Adam.

Adam is key for faster training. It ensures stable training. It is essential for deep learning.

- Experiment with Adam in your networks.
- Learn about advanced optimizers.