



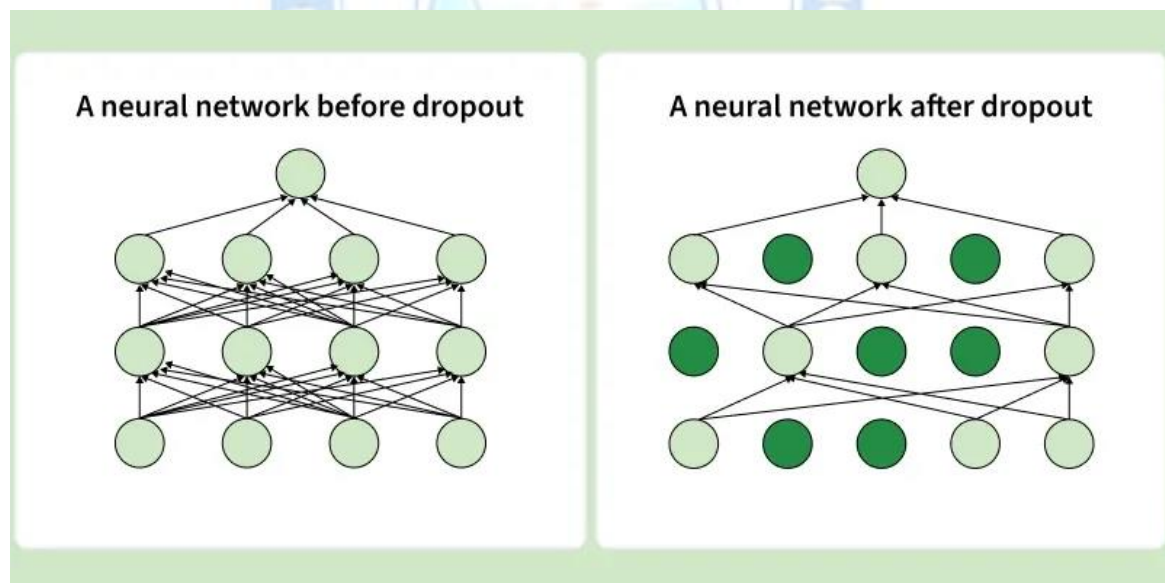
# ROHINI

## COLLEGE OF ENGINEERING & TECHNOLOGY

Approved by AICTE and Affiliated to Anna University (An ISO Certified Institution) | Accredited with A+ Grade by NAAC  
Recognized under Section 2(f) of University Grants Commission, UGC ACT 1956  
**(AUTONOMOUS)**

### DROPOUT

Training deep learning models for too long on the same data can lead to overfitting where the model performs well on training data but poorly on unseen data. Dropout Regularization helps overcome this by randomly deactivating a portion of neurons during training hence forcing the model to learn more robust and independent features.



This image shows how dropout works by randomly deactivating neurons during training, forcing the network to rely on multiple paths and preventing overfitting

### Why Dropout is needed

Dropout is needed because it prevents the network from relying too much on specific neurons, hence improving generalisation by training and reducing co-adaptation among neurons. Below are the key features of dropout:

- Randomly deactivates a set percentage of neurons in each training step.
- Works with dense, convolutional and recurrent layers.
- Dropout rate controls how many neurons are dropped.
- Remaining active neurons are scaled to maintain output stability.
- Reduces co-dependency among neurons and improves generalization.

### How it Works

Dropout works by randomly turning off a fraction of neurons during each training pass. When a neuron is dropped it sends no output and receives no weight updates for that batch. The dropout rate  $P$  determines the probability of removing a neuron. A neuron output  $h$  is modified as

$$h' = h \cdot m$$

where

- ✓  $m$ : Dropout mask randomly sampled for each neuron.
- ✓  $m = 1$ : neuron is kept
- ✓  $m = 0$ : neuron is dropped

To keep activation values consistent during training inverted dropout is used where active neurons are scaled as:

$$h' = \frac{h \cdot m}{1 - P}$$

where

- ✓  $h$ : Original output of a neuron before dropout.
- ✓  $h'$ : Output after applying dropout (scaled or zeroed out).

This ensures that the expected output remains stable even when some neurons are dropped.

*Note: During testing dropout is disabled meaning all neurons remain active and no randomness is applied. The network uses its full capacity and activations are scaled appropriately to match the expected values from training.*

### **Types of Dropout Regularization**

- **Standard Dropout:** Randomly removes individual neurons during training to reduce overfitting.
- **Spatial Dropout:** Drops entire feature maps in CNNs to preserve spatial structure.
- **Dropout2D / Dropout3D:** Applies dropout across 2D or 3D feature channels for images and volumetric data.
- **DropConnect:** Randomly drops weights instead of neurons, offering stronger regularization.
- **Alpha Dropout:** Maintains activation statistics and works best with SELU-based networks.
- **Variational Dropout:** Uses the same dropout mask across time steps in RNNs for stable learning.
- **Monte Carlo Dropout:** Keeps dropout active at inference to estimate prediction uncertainty.

