AI TERMINOLOGY / LEVEL 1 / GROUP 1

Deep Learning

From Basics to Advanced Research

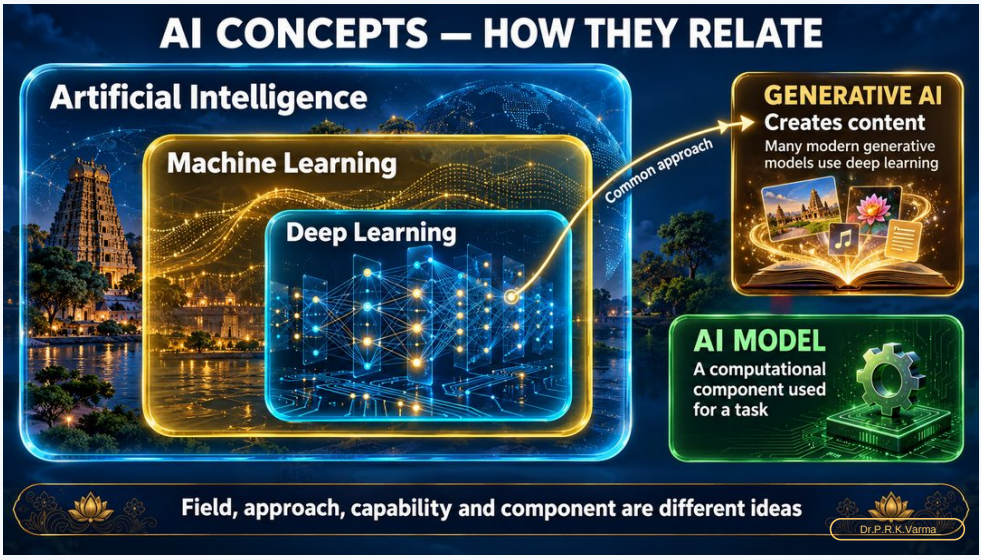


Presenter guide with slide visuals and notes. The slide numbers match the PowerPoint.

BASICS 02–25

ADVANCED 26–36

RESEARCH 37–47



AI Concepts: How They Relate

PRESENTER NOTES

Machine learning is part of AI, and deep learning is part of machine learning. Generative AI describes content creation, while an AI model is a working computational component, not another nested field.



A Learnable Detector

PRESENTER NOTES

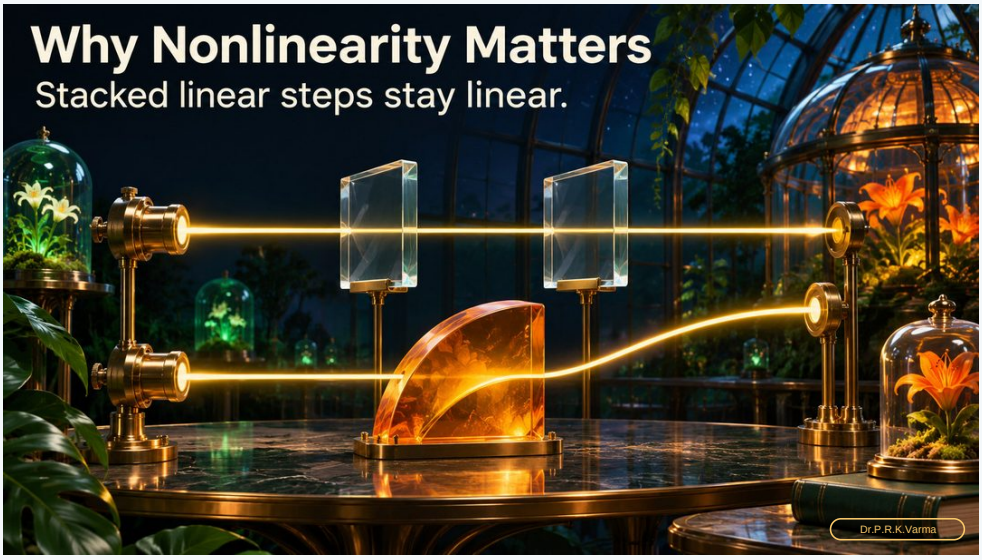
A neuron combines input signals using learned weights and applies a response rule. Some units become sensitive to a useful pattern, but their role is learned and may not be easy to name.



Layers and Depth

PRESENTER NOTES

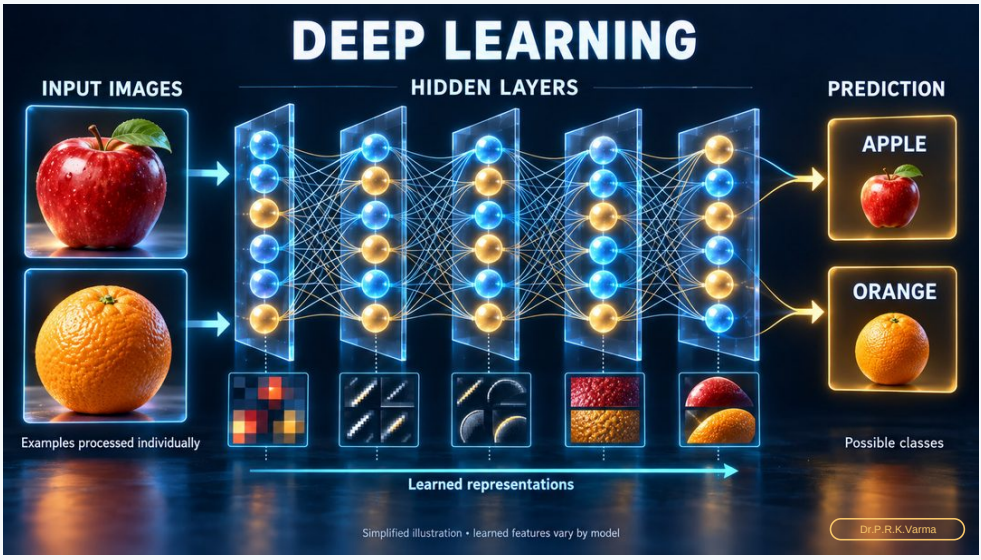
Layers transform one representation into the next. Depth refers to how many successive transformations are used; it can support more complex representations, but depth alone does not guarantee useful features.



Why Nonlinearity Matters

PRESENTER NOTES

Stacking two linear transformations gives another linear transformation. Nonlinear activations make it possible for a network to represent relationships that a single linear mapping cannot. The bent beam is a metaphor, not an activation graph.



Deep Learning: Fruit through Layers

PRESENTER NOTES

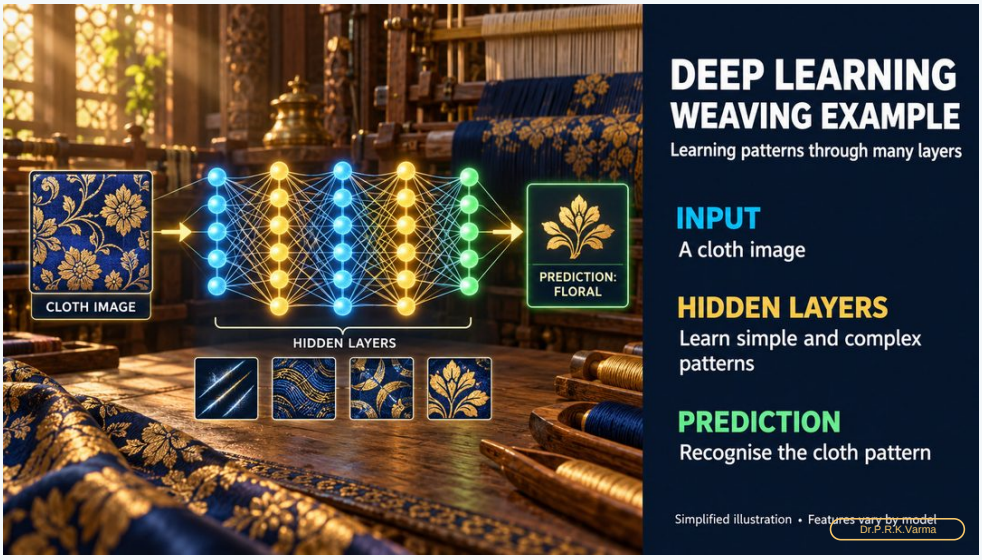
An apple or orange image passes through a layered neural network to produce a prediction. The diagram simplifies learned representations; five displayed layers do not mean all networks have five layers or fixed roles for each layer.



Layered Recognition of Rangoli

PRESENTER NOTES

A rangoli image is transformed through learned layers before a category is predicted. The illustrated progression from simple to complex features is a teaching aid, not a guaranteed map of what every layer learns.



Layered Recognition of Cloth

PRESENTER NOTES

A neural network can learn useful representations of cloth images for pattern classification. Many layers help transform inputs, but the colourful layer icons do not establish the behaviour of an actual trained system.



Tensors and Channels

PRESENTER NOTES

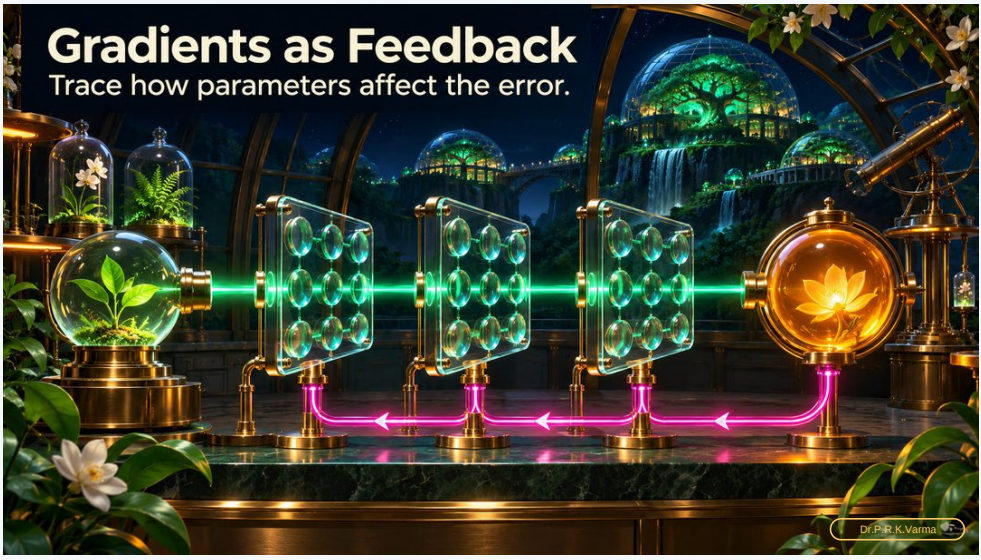
A tensor stores values along organized dimensions. An image tensor can preserve height, width, and channels such as red, green, and blue. The channel order matters when sending a tensor to a model.



The Forward Pass

PRESENTER NOTES

The forward pass computes each layer's output in sequence and produces a prediction. During training the prediction is used to calculate a loss. The forward pass alone does not update weights.



Gradients as Feedback

PRESENTER NOTES

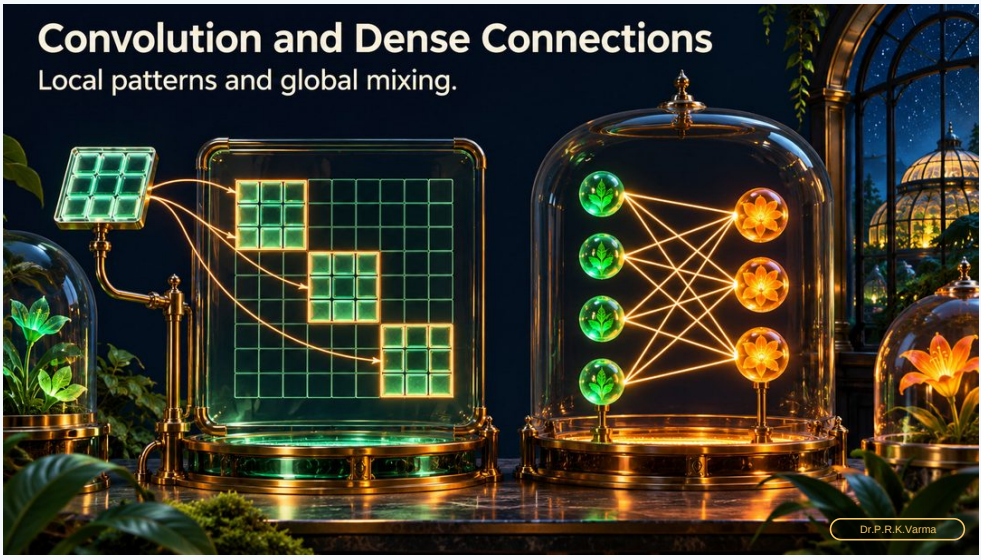
After loss is computed, backpropagation calculates how small changes to parameters would affect it. The optimizer then uses those gradients to choose updates. Gradients are not additional labels.



Receptive Fields

PRESENTER NOTES

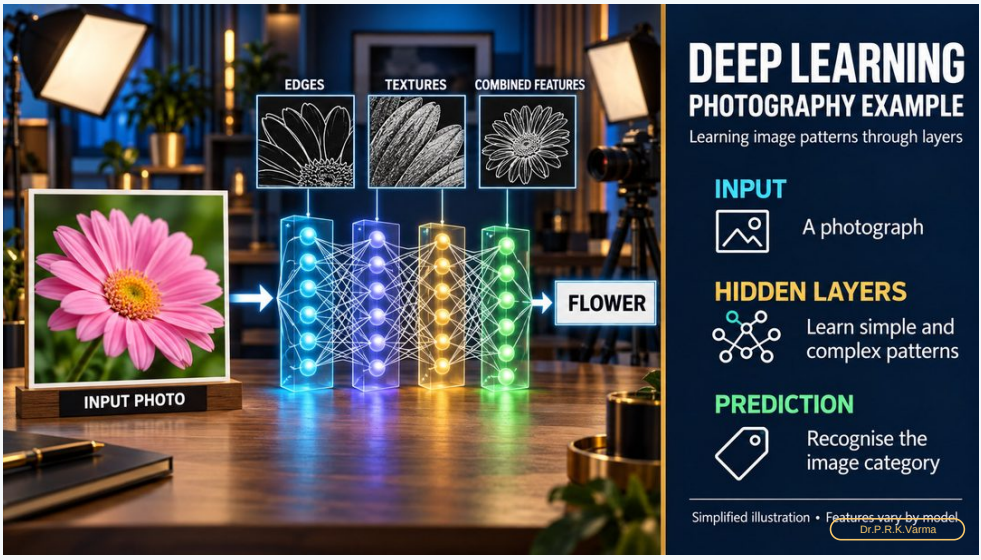
The receptive field of a unit is the region of the original input that can affect it. Stacking many local operations often expands this region, although not all positions necessarily influence the result equally.



Convolution and Dense Connections

PRESENTER NOTES

Convolutions commonly apply shared weights to local regions. A dense layer connects every output unit to every input of that layer; it does not impose the same local pattern-sharing structure.



Edges, Textures and Combined Features

PRESENTER NOTES

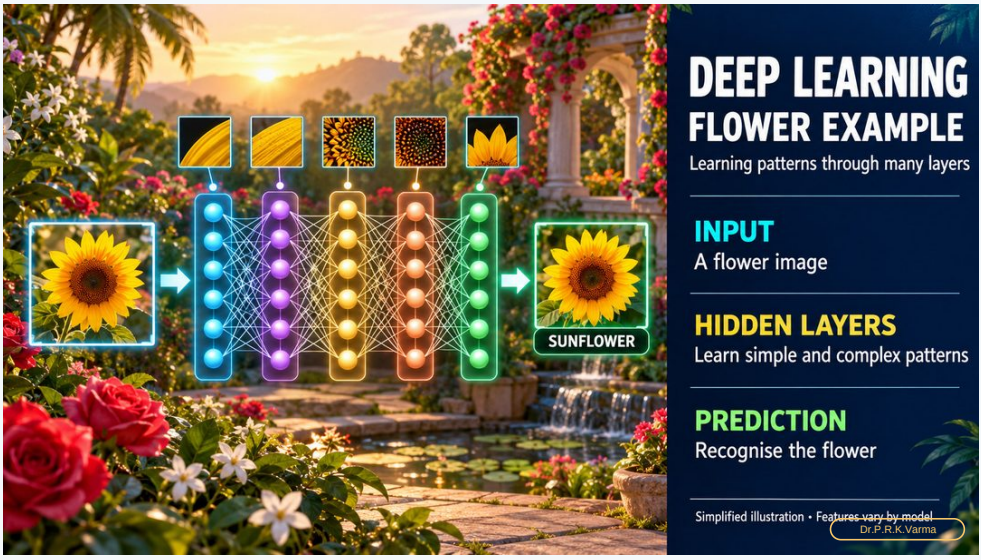
The photography example connects image features to a category prediction. Edges, textures and combinations are useful visual intuitions, but exact representations and layer roles vary by model and task.



Deep Learning for Sound Patterns

PRESENTER NOTES

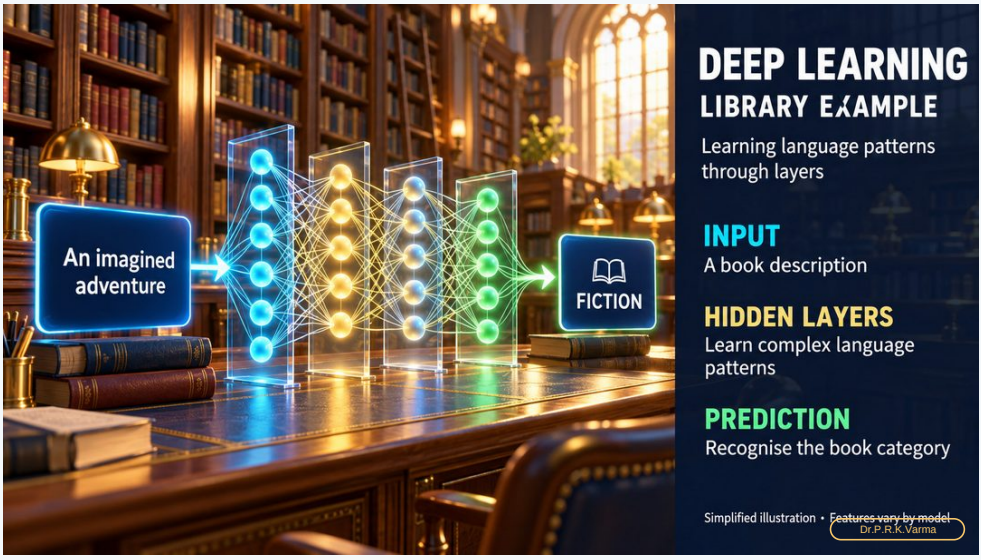
A sound recording passes through learned layers before an instrument prediction. The example illustrates representation learning for audio; it does not claim a human-like understanding of music.



Deep Learning for Flower Images

PRESENTER NOTES

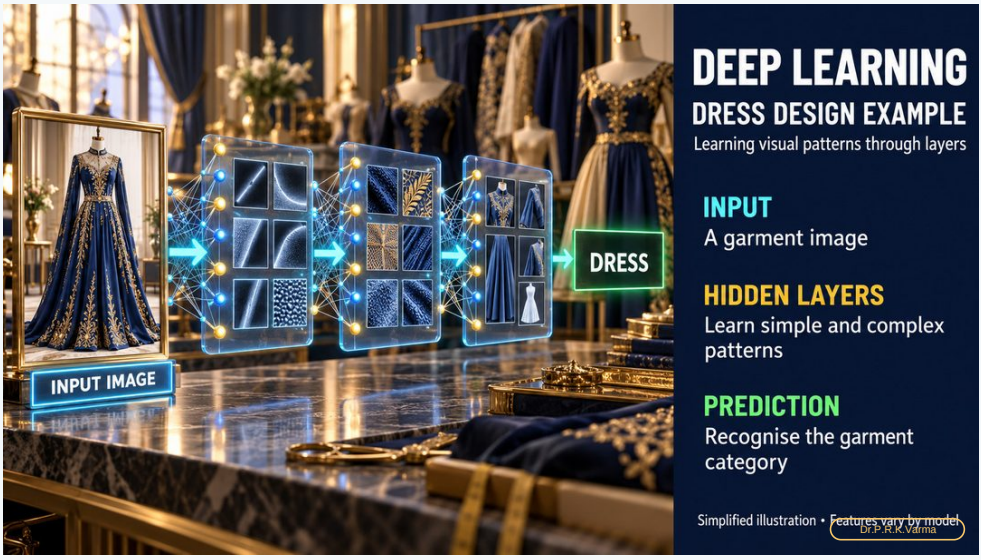
Learned layers transform a flower image into features useful for classification. The sunflower output is an illustrative prediction, not evidence that this pictured network has been trained or tested.



Deep Learning for Language Patterns

PRESENTER NOTES

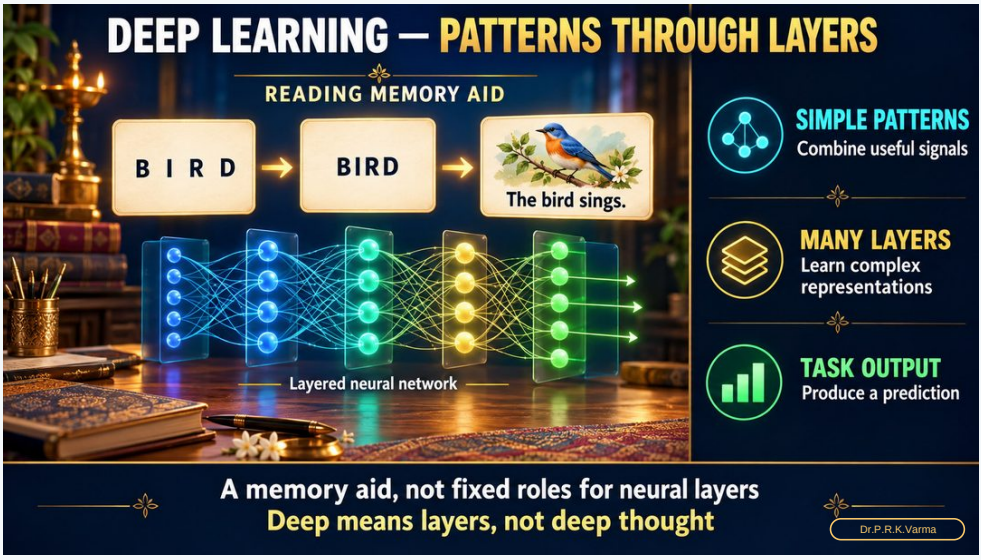
A book description is processed through layers to predict a category such as fiction. Language patterns can be learned computationally, while the illustrated network remains a simplified teaching diagram.



Deep Learning for Garment Images

PRESENTER NOTES

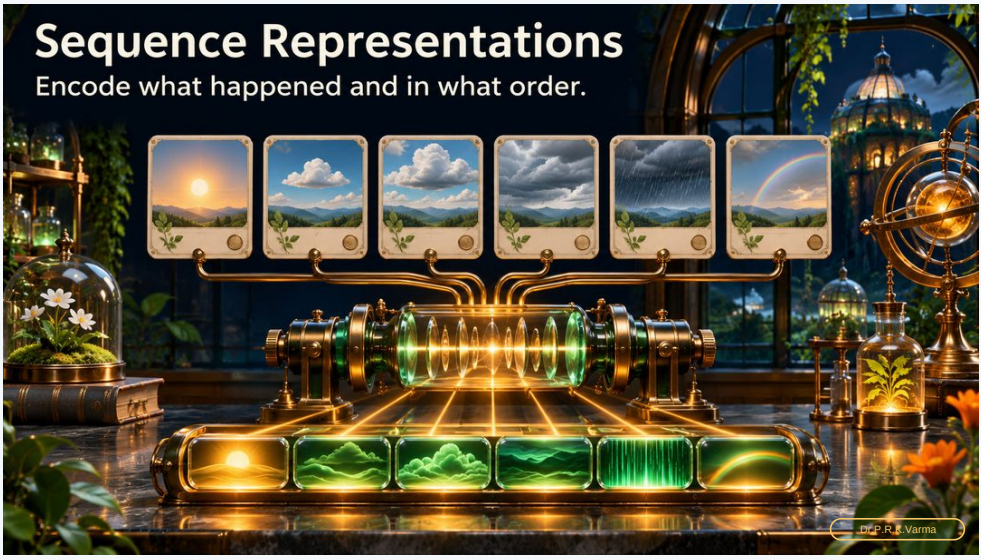
A layered network learns visual patterns relevant to garment categories. Recognition is the task shown here, not creating a clothing design, and the displayed prediction needs real evaluation before practical use.



Letters, Words and Sentences: A Memory Aid

PRESENTER NOTES

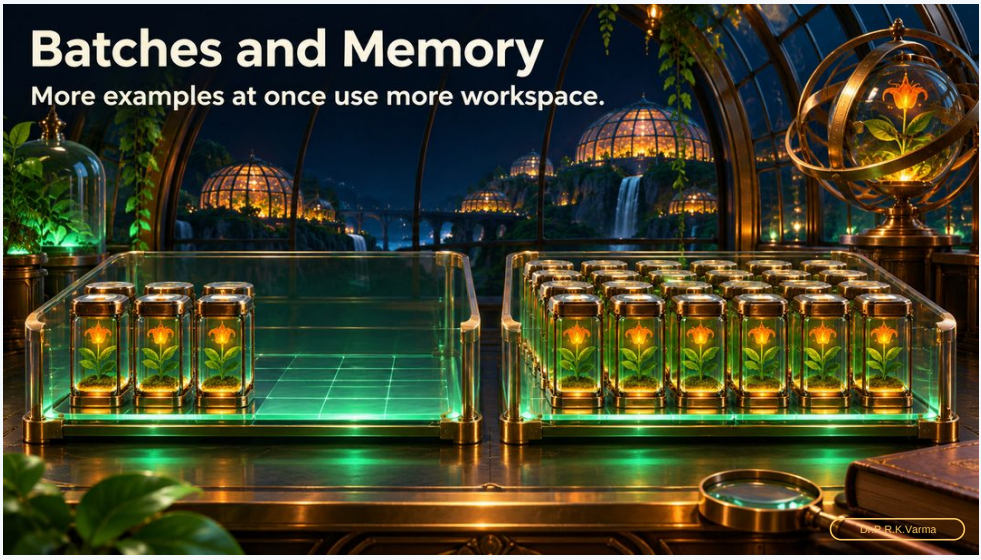
Combining small elements into richer patterns offers an intuition for layered representations. Neural layers are not literally assigned one each to letters, words and sentences. Deep refers to many layers, not deep thought.



Sequence Representations

PRESENTER NOTES

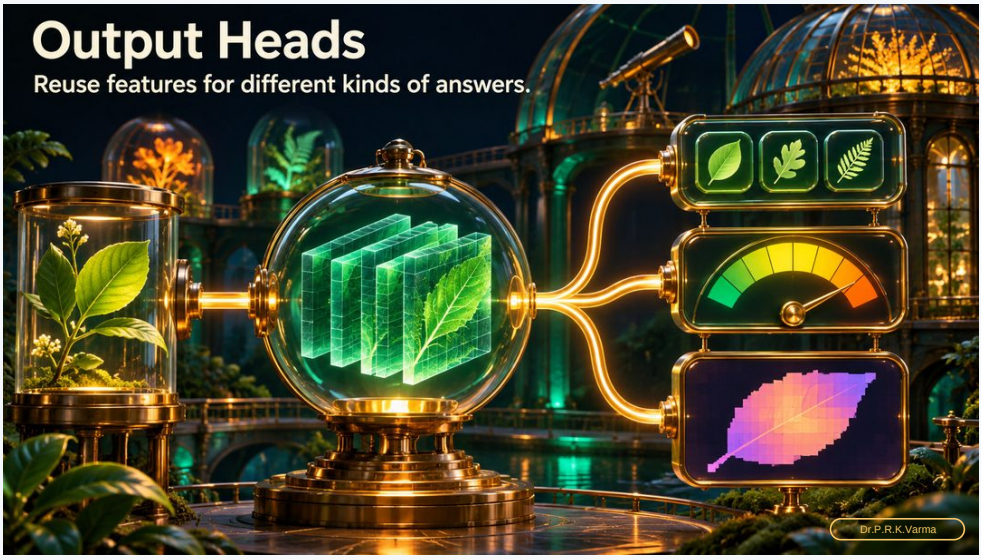
A sequence representation preserves information about position or order, such as the timing of sensor readings or word order in a sentence. Different architectures handle that ordering in different ways.



Batches and Memory

PRESENTER NOTES

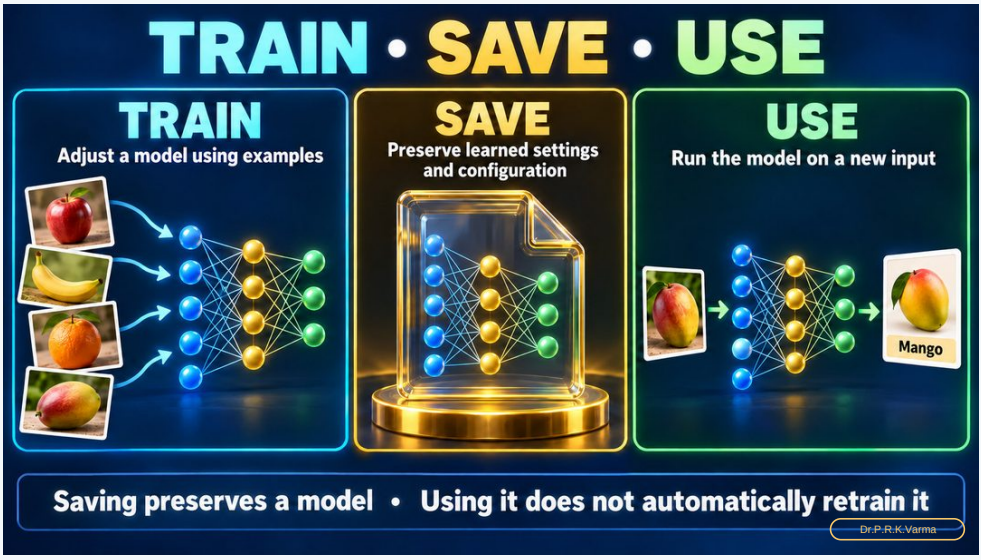
A batch is a set of examples processed together. Larger batches often increase memory needed for data and intermediate values. Fit the batch size to the hardware and task rather than assuming bigger is better.



Output Heads

PRESENTER NOTES

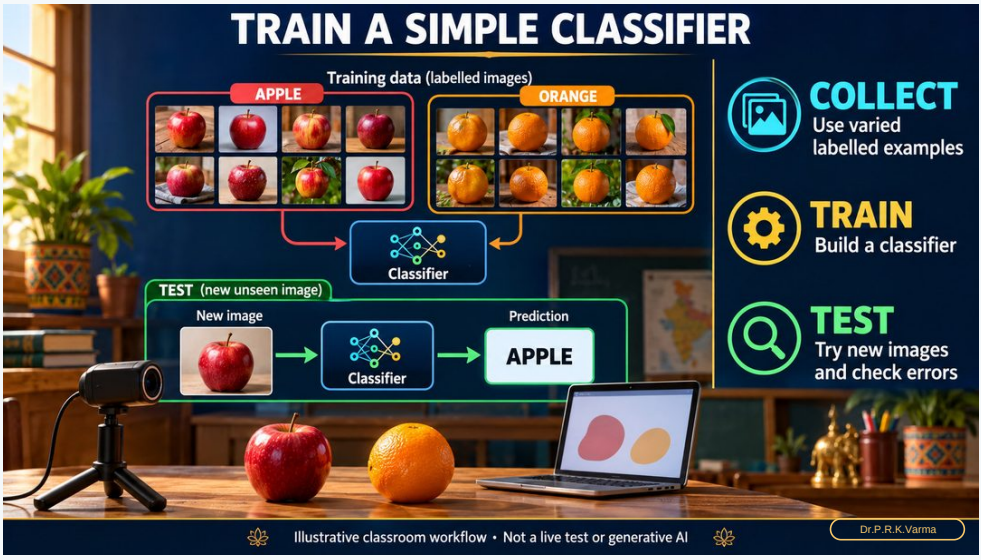
An output head turns shared features into the required prediction. A model may use different heads for classification, numerical estimates, or pixel-level labels. Each head has its own output format and learning target.



Train, Save, Use

PRESENTER NOTES

Training adjusts a model using examples. Saving preserves learned settings, and use applies the model to a new input. Saving does not turn an output into a model, and ordinary use does not retrain it.



A Practical Classification Demonstration

PRESENTER NOTES

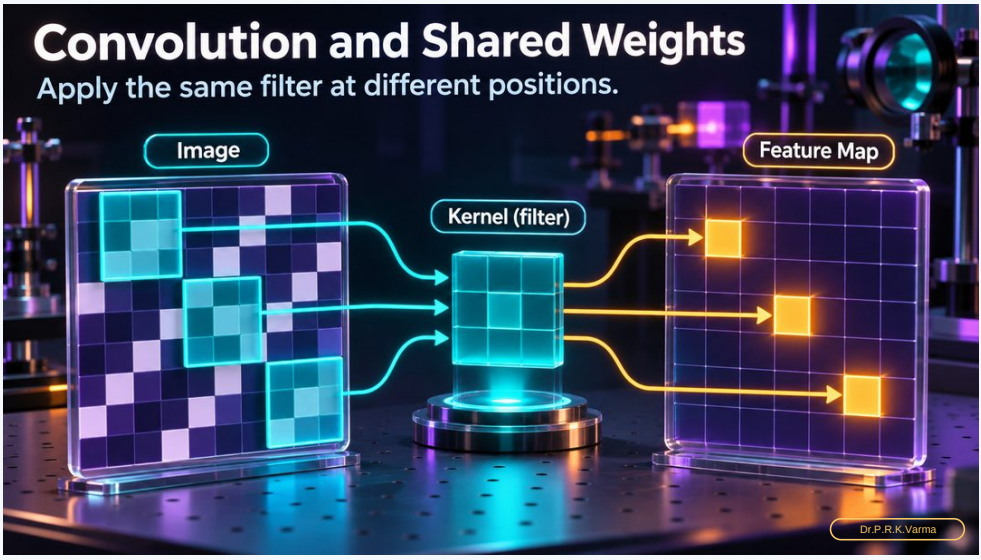
Train with labelled apple and orange images, then classify a new image. This is a teaching illustration, not the measured result of an experiment that was actually run.



Five Ms: A Memory Aid

PRESENTER NOTES

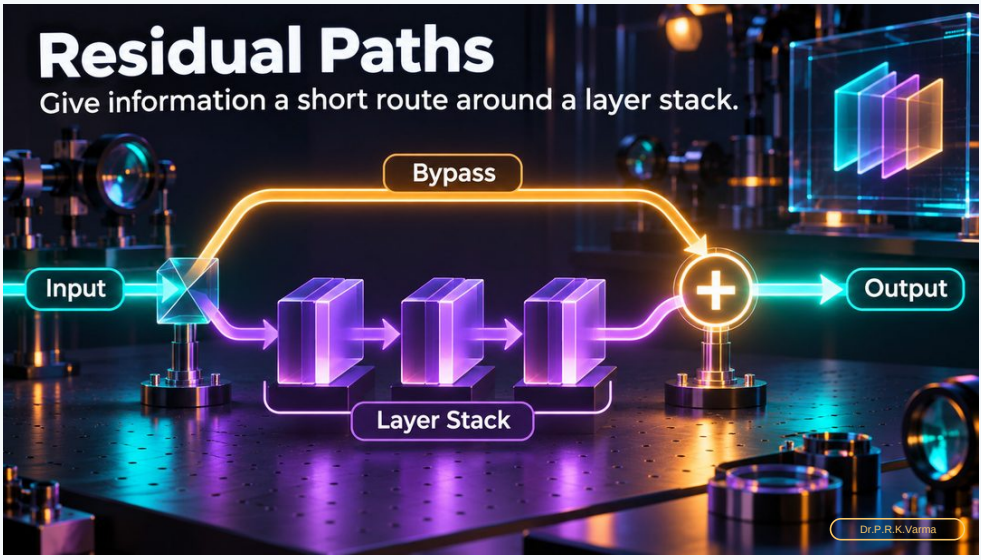
Machines, Many Examples, Many Layers, Making Content and Model for Reuse recall five related concepts. They are memory hooks, not five training stages or a chain of nested categories.



Convolution and Shared Weights

PRESENTER NOTES

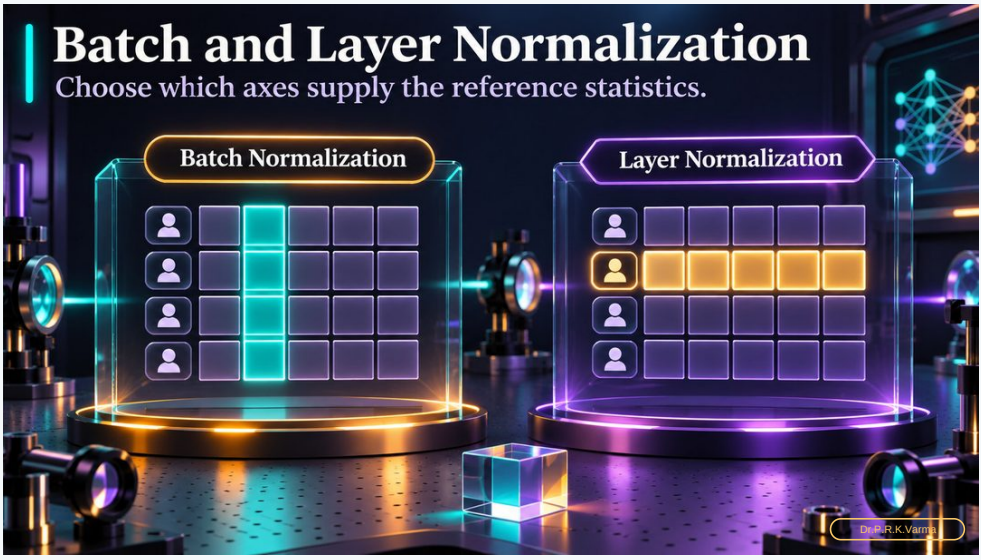
A convolution applies a shared learned kernel across spatial positions. Reuse reduces parameter count and lets a detector respond at multiple locations, subject to stride, padding and boundary choices.



Residual Paths

PRESENTER NOTES

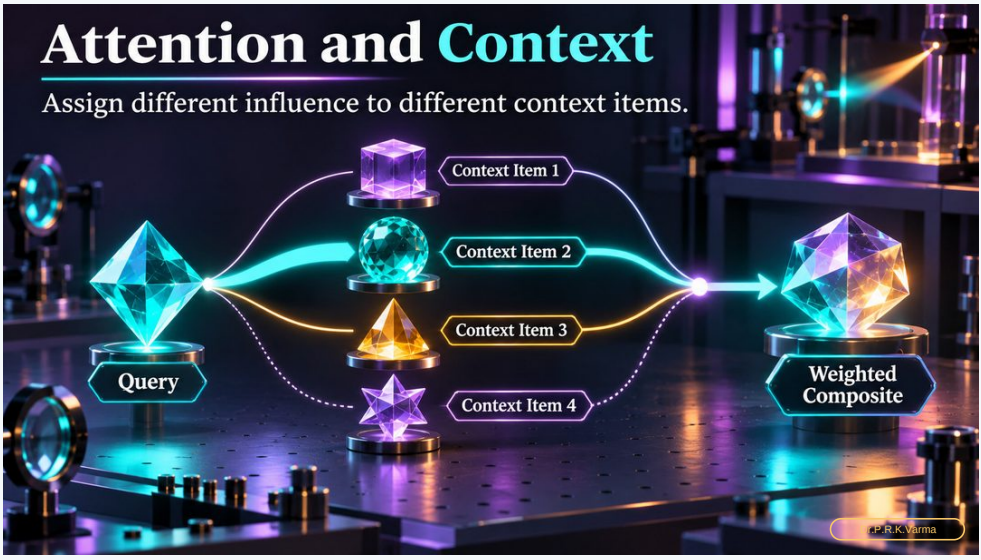
A residual block combines a learned transformation with a skip path. The shortcut helps information and gradients propagate; a projection can align dimensions if the block changes shape.



Batch and Layer Normalization

PRESENTER NOTES

Batch normalization uses batch-derived statistics during training, usually with running estimates at inference. Layer normalization operates across the chosen features of each example. Select based on architecture and workload.



Attention and Context

PRESENTER NOTES

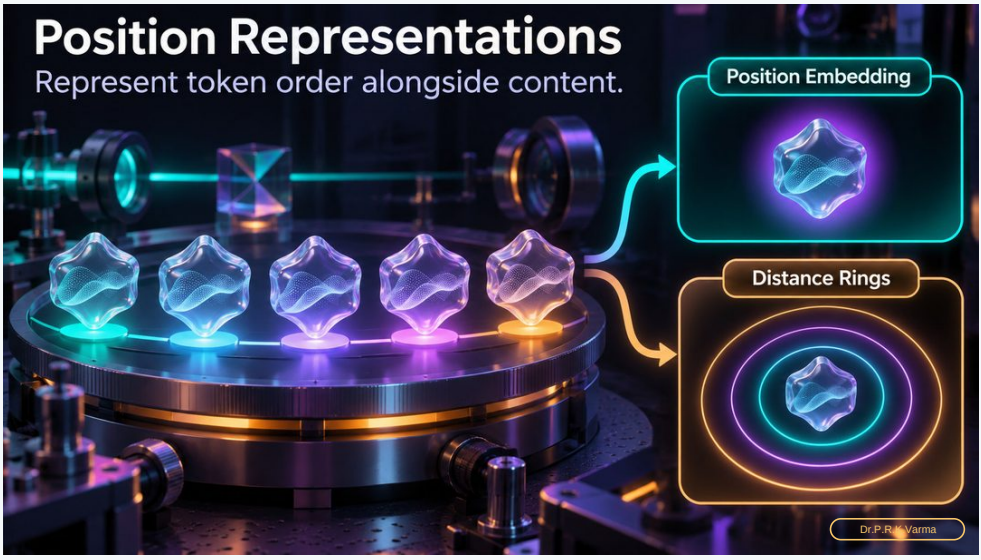
An attention layer scores the match between a query and keys, then uses those weights to combine values. Distinct heads can use different projections; weight patterns are not automatically explanations.



Causal Masking

PRESENTER NOTES

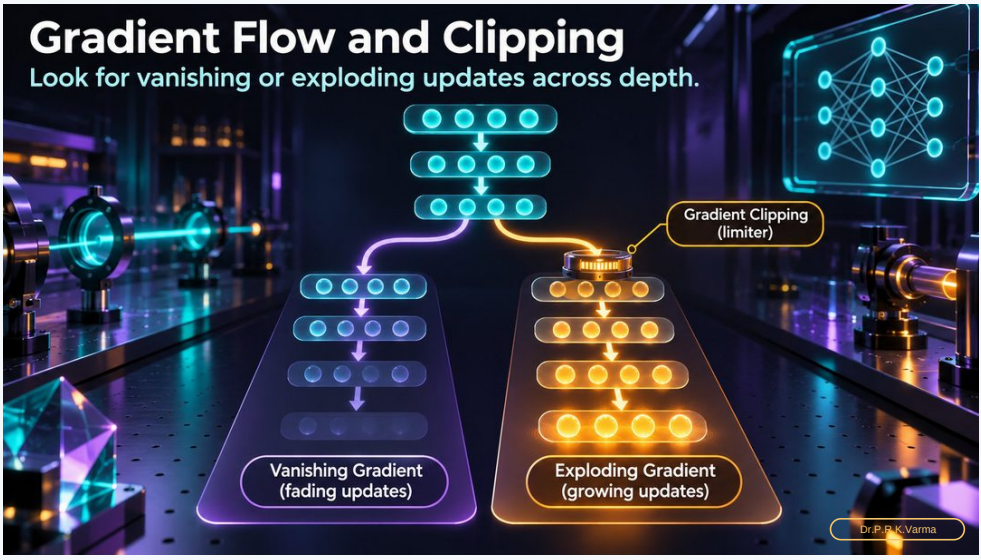
In autoregressive training, a causal mask prevents a position from attending to future tokens. Check the mask and target alignment so the model cannot read the answer while training.



Position Representations

PRESENTER NOTES

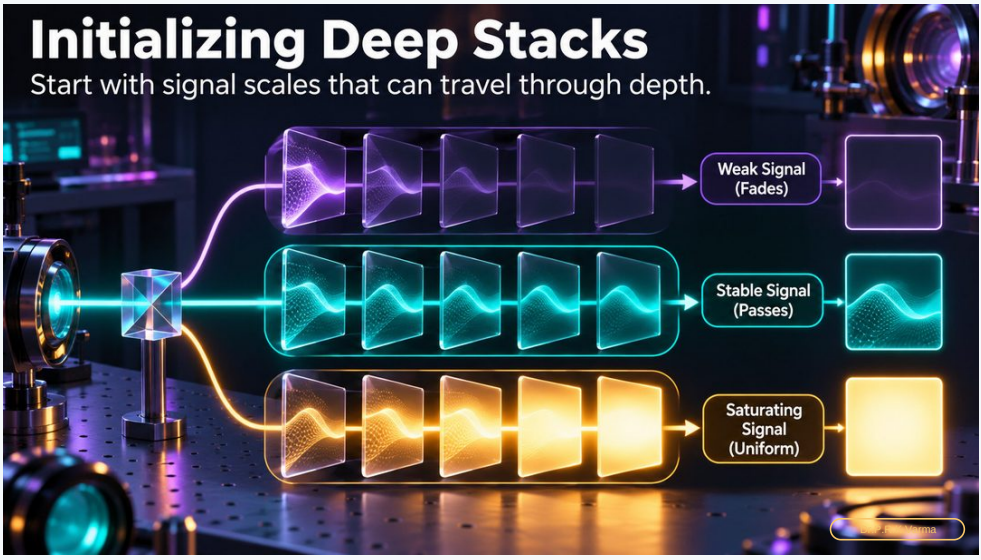
Attention without position information cannot distinguish permutations through token order alone. Absolute embeddings attach positions to tokens; relative or rotary approaches introduce positional relationships differently.



Gradient Flow and Clipping

PRESENTER NOTES

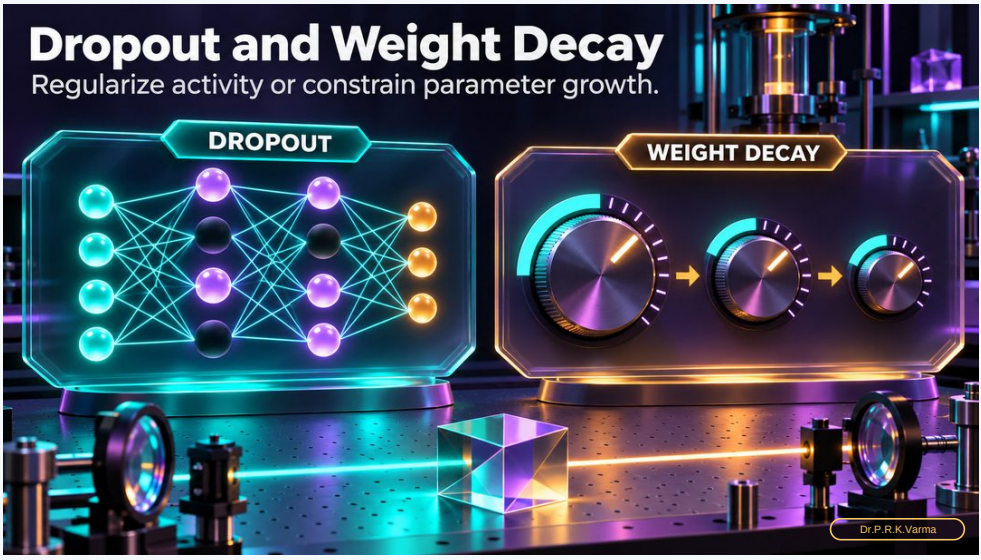
Inspect layerwise gradient norms and nonfinite values during training. Global-norm clipping caps large updates, not small ones; vanishing gradients need changes such as initialization, architecture or optimization.



Initializing Deep Stacks

PRESENTER NOTES

Choose initialization suited to the activation and architecture. Check activation and gradient statistics through depth on a small batch; even principled initialization does not guarantee stable training.



Dropout and Weight Decay

PRESENTER NOTES

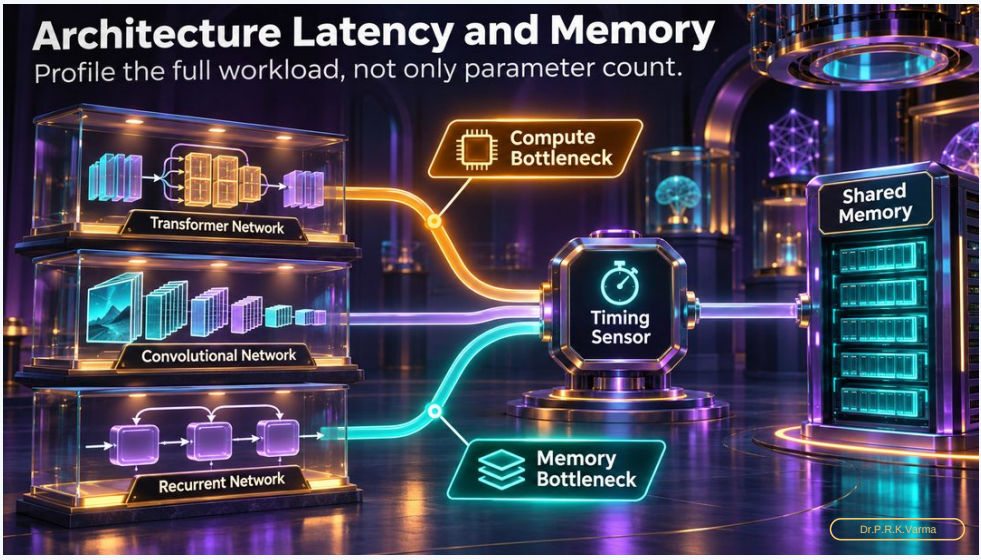
Dropout masks activations at training time and is typically disabled at inference. Weight decay shrinks parameters during optimization; decoupled weight decay is not generally equivalent to an L2 penalty under adaptive optimizers.



Activation Checkpointing

PRESENTER NOTES

Save only selected forward activations and recompute others during the backward pass. Account for additional computation, randomness and memory peaks when comparing checkpointing strategies.



Architecture Latency and Memory

PRESENTER NOTES

Latency depends on operation shapes, hardware, memory movement, batch size and implementation. Profile peak memory, cold-start and steady-state latency against representative inputs before deciding on an architecture.



Effective Receptive Fields

PRESENTER NOTES

The theoretical field describes all input positions that can influence a unit. The effective field measures which positions matter most under a specified measurement and may occupy a smaller region.

Original research

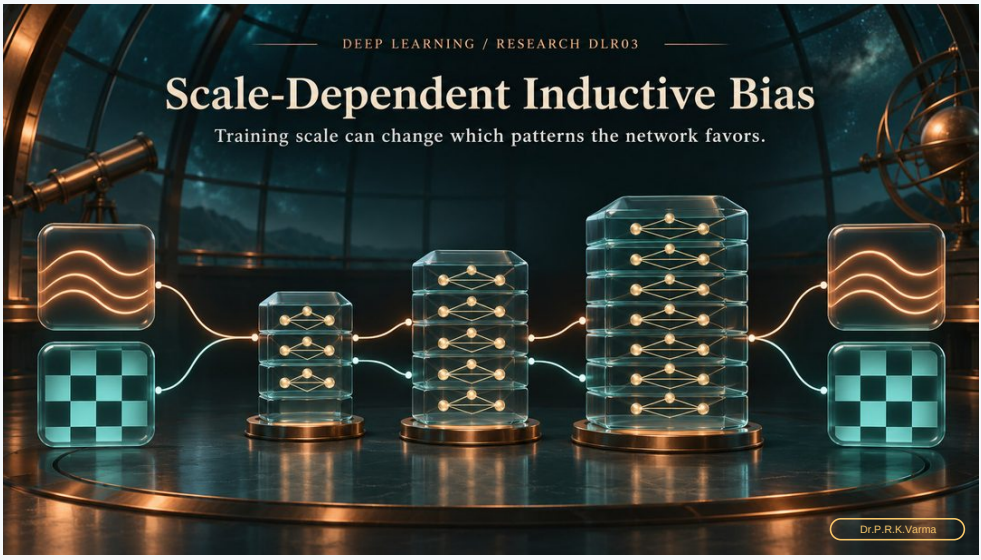


Residual Iterative Refinement

PRESENTER NOTES

One interpretation of residual layers is iterative estimation: a block adjusts an existing representation through a residual branch. This is a research interpretation, not a requirement that every block refines the same semantic quantity.

Original research

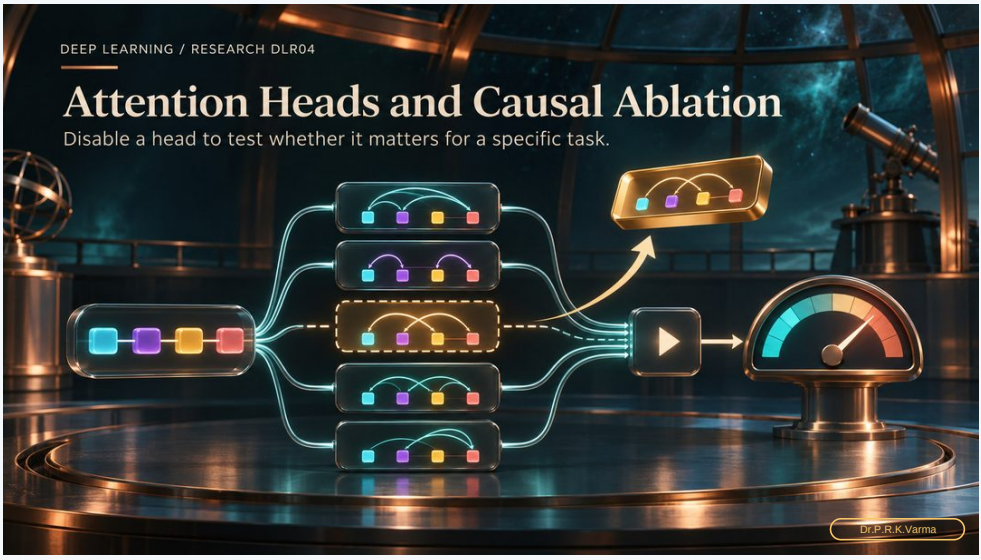


Scale-Dependent Inductive Bias

PRESENTER NOTES

Vary model and training scale while holding the data-generating setup and evaluation protocol fixed. Results such as deep double descent show scale-dependent changes in generalization in studied settings, not a universal ranking or mechanism.

Original research

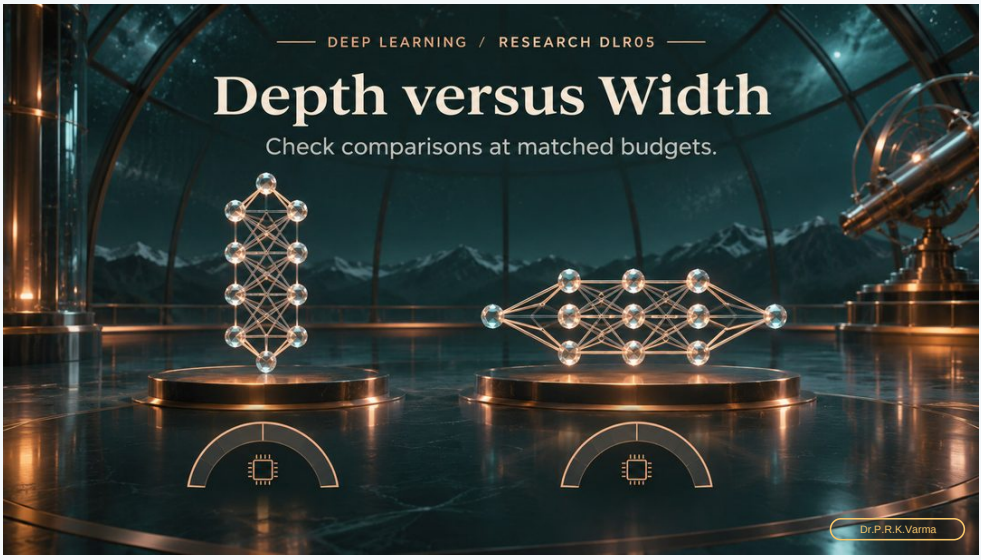


Attention Heads and Causal Ablation

PRESENTER NOTES

Remove or disable one head, then compare measured performance against a baseline and control. An observed change indicates a role for that intervention and task; individual importance can differ from joint importance.

Original research

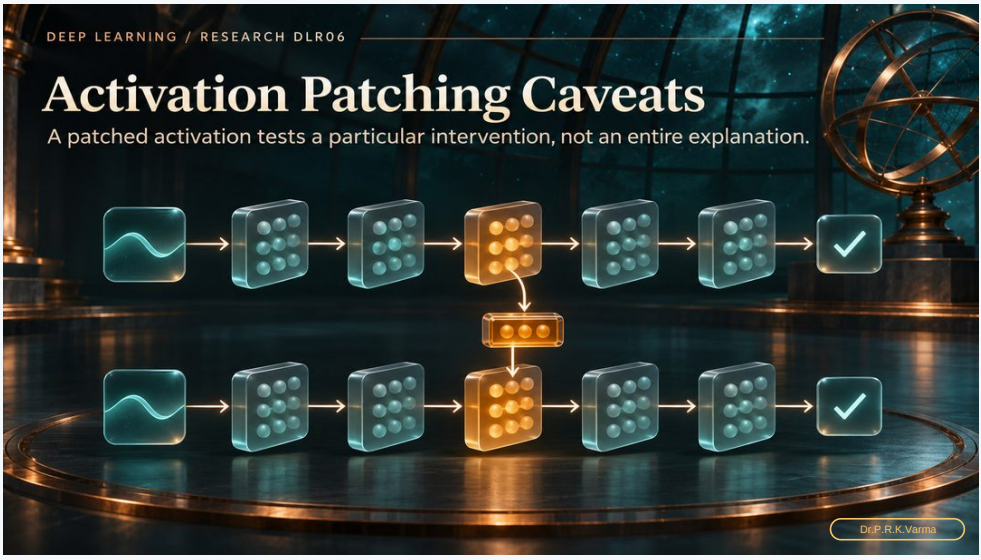


Depth versus Width

PRESENTER NOTES

Shallower, wider residual architectures have worked well in studied settings. Compare depth and width at matched compute or report the unmatched budgets explicitly; also measure latency and performance across seeds.

Original research



Activation Patching Caveats

PRESENTER NOTES

Replace a particular activation with one from a paired input and measure the change. Outcomes can depend on the baseline, intervention site and metric, so validate with alternatives before making a broad explanatory claim.

Original research



Feature Superposition

PRESENTER NOTES

In toy models, sparse features can share a smaller number of representational dimensions while interfering with one another.
Similar-looking patterns in a larger model require additional evidence before identifying specific features.

Original research



Grokking in Controlled Tasks

PRESENTER NOTES

Delayed generalization has been documented in small algorithmic tasks under particular settings. Compare held-out curves across seeds and training lengths; a late improvement is not guaranteed in other tasks.

Original research



Sparsity and Pruning Trade-offs

PRESENTER NOTES

Pruning reduces the number of nonzero weights, but runtime gains depend on sparse-kernel support, memory formats and hardware. Check the resulting quality, model footprint and measured latency separately.

Original research

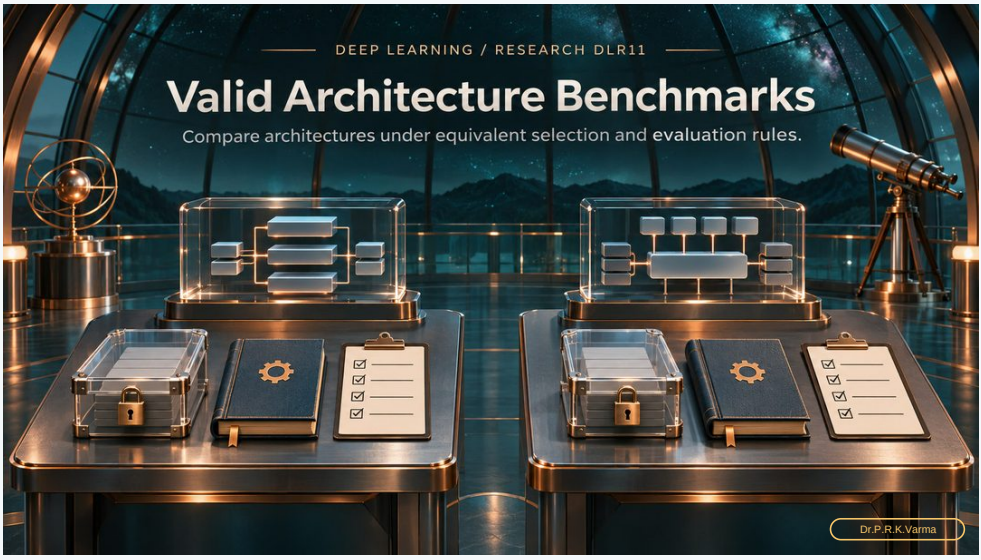


Causal Controls for Interpretability

PRESENTER NOTES

A proposed mechanism should make predictions about targeted interventions and controls. Test alternatives and random or matched replacements to see whether the evidence distinguishes the explanation from competing stories.

Original research



Valid Architecture Benchmarks

PRESENTER NOTES

Match data, evaluation splits, selection rules and budgets where possible, and report differences explicitly. Training-system studies show that conclusions can vary with workload, optimization settings and resource definitions.

Original research