
🟦 Fundamental Math & Learning Concepts

- Gradient
- Derivative
- Partial derivative
- Chain rule
- Loss
- Logarithm (log base e / natural log / ln)
- Weighted sum
- Dot product
- Probability vector
- Exponentiation / exp()
- Logits
- One-hot vector

🟥 Training Stability & Optimization

- Vanishing gradient
- Exploding gradient
- Backpropagation
- Optimization
- Learning (via gradient updates)
- Loss function
- Cross-entropy loss
- Training convergence
- Smooth loss
- Stability
- Update (in learning)

🟧 Neural Network Components

- Neuron
- Dense layer
- Fully connected layer
- Hidden layer

- Input layer
- Output layer
- Activation function
- Softmax
- Argmax (mentioned indirectly in softmax)
- ReLU (ReLU mentioned earlier)
- Dropout (mentioned in code)

🟩 CNN / Image-specific

- Convolutional layer
- Filter
- Feature map
- Channel
- Padding
- Stride
- GlobalAveragePooling2D
- Batch size
- Batch
- Batch normalization
- Cross-correlation
- Convolution
- Residual connection
- Shortcut (residual block)

🟪 Numerical Methods & Frameworks

- Taylor series
- Polynomial approximation
- Padé approximant
- Piecewise function
- Softmax implementation
- tf.exp()
- tf.nn.softmax()
- Weight matrix
- Bias vector

🟫 Libraries & Hardware Concepts

- TensorFlow
- CUDA
- Eigen
- Hardware-accelerated math
