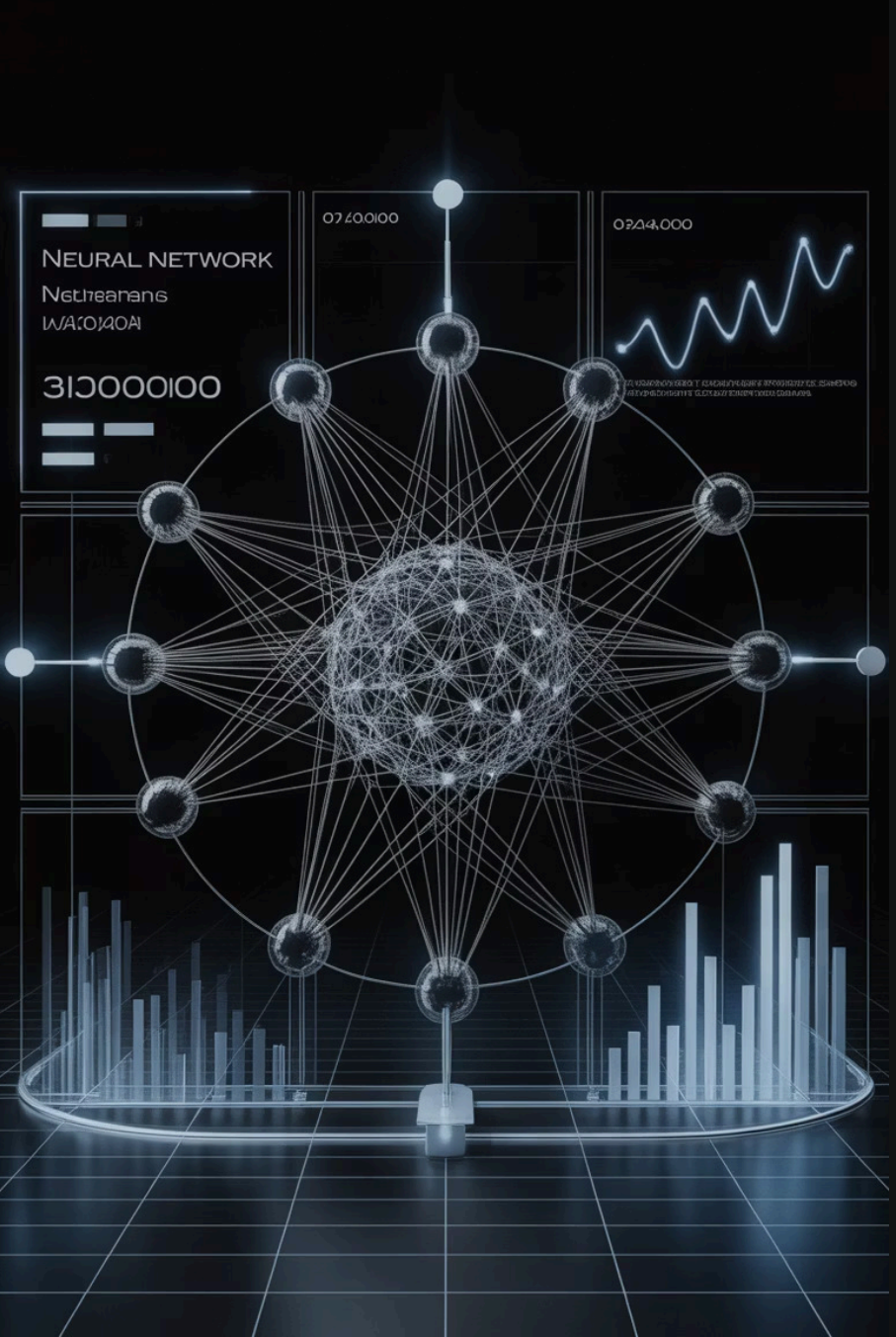


Evaluating & Improving Deep Learning Models

Welcome to Week 4! This week we'll explore how to properly evaluate deep learning models and implement techniques to enhance their performance.

by S MM

Understanding Evaluation Metrics

1. **Accuracy**

   The proportion of correct predictions among all predictions made.

2. **Loss**

   Measures the difference between predicted and actual values.

3. **Precision**

   The ratio of true positives to all positive predictions.

4. **Recall**

   The ratio of true positives to all actual positives.

5. **F1-Score**

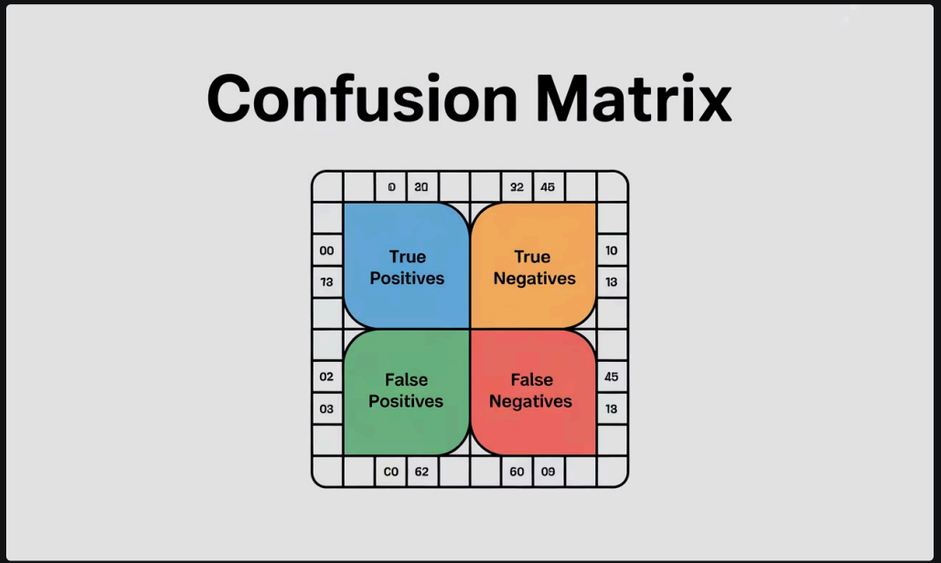
   The harmonic mean of precision and recall.

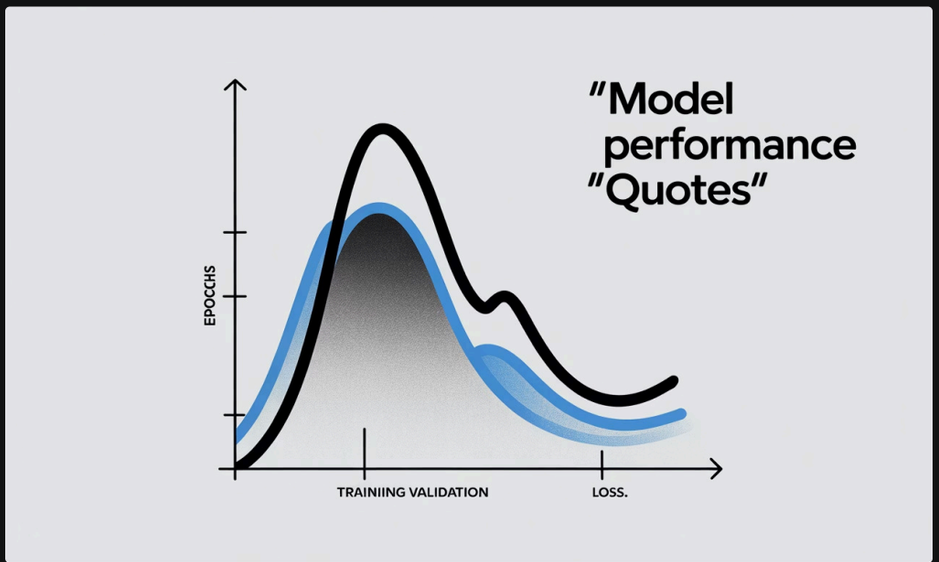
Visualizing Model Performance

Confusion Matrix

A table showing true positives, false positives, true negatives, and false negatives.

Helps identify which classes your model struggles with most.

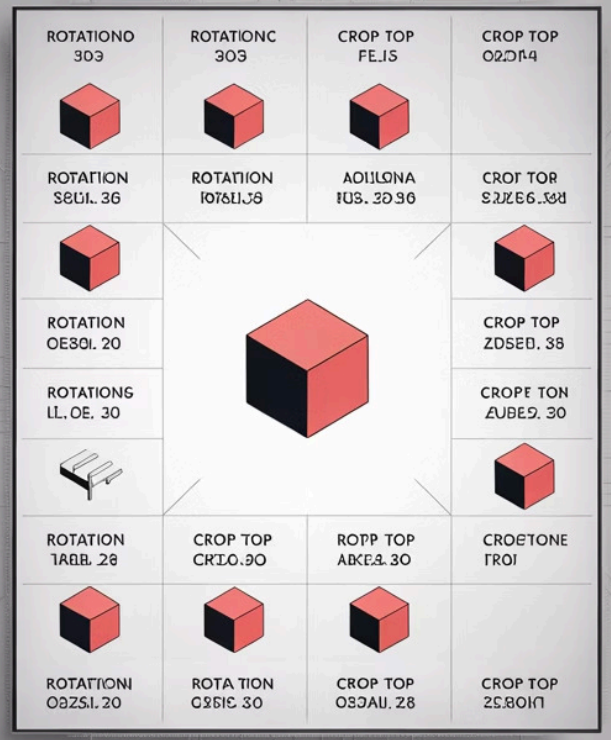




Learning Curves

Plots showing how training and validation metrics change over time.

Helps identify overfitting, underfitting, and convergence patterns.



Data Augmentation

Original Data

Limited training examples that may not cover all variations.

Transformation

Apply flips, rotations, crops, color shifts, and noise.

Expanded Dataset

More diverse training examples without collecting new data.

Improved Generalization

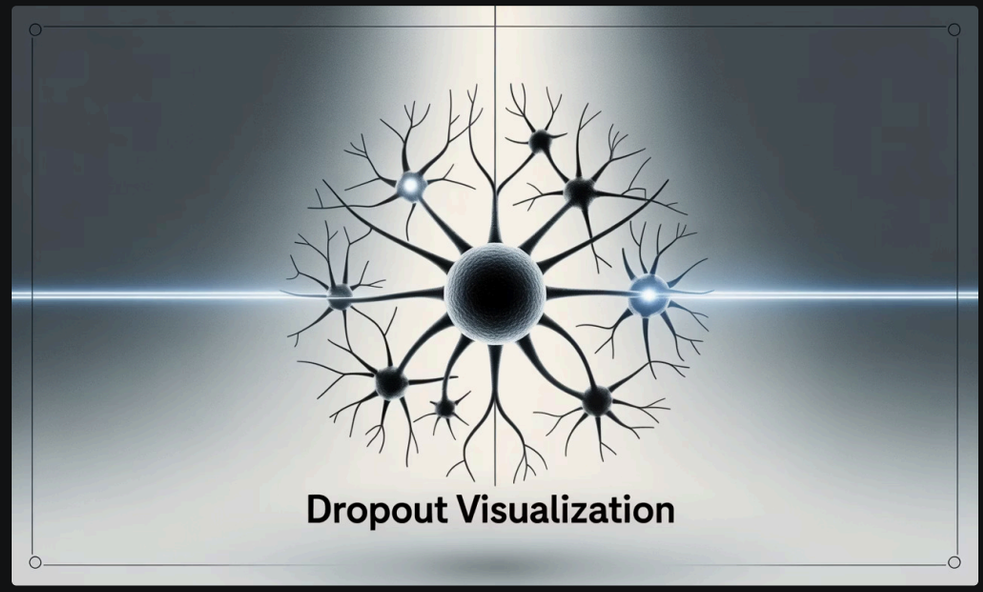
Model learns invariant features rather than memorizing examples.

Dropout & Regularization

Dropout Technique

Randomly deactivates neurons during training to prevent co-adaptation.

- Typically set between 0.2 and 0.5
- Acts as ensemble learning
- Reduces overfitting significantly



1

Without Dropout

Network memorizes training data, performs poorly on new examples.

2

With Dropout

Network learns robust features, generalizes better to unseen data.

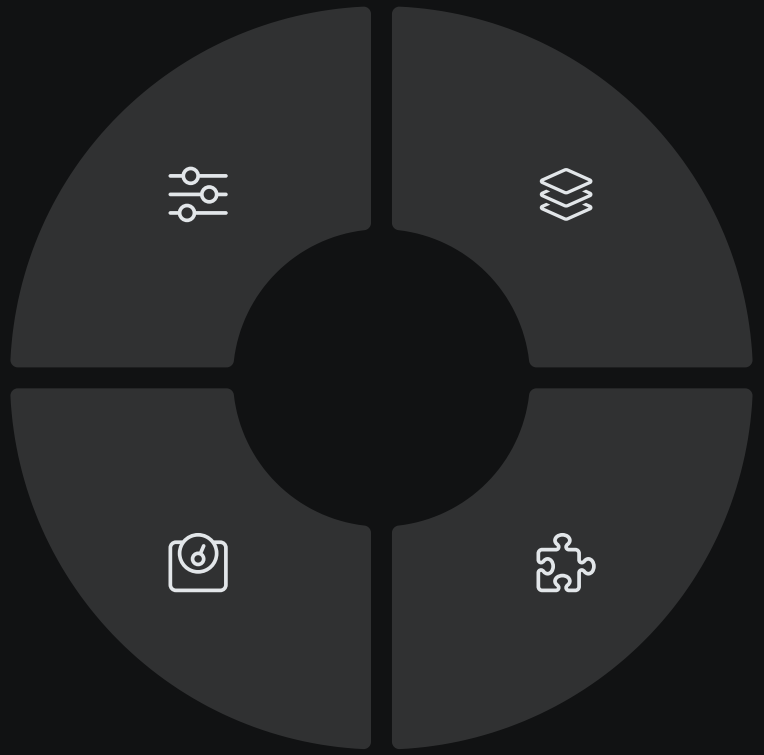
Hyperparameter Tuning

Learning Rate

Controls step size during optimization. Too high: unstable. Too low: slow convergence.

Regularization

L1/L2 penalties prevent excessive weight values and improve generalization.



Network Architecture

Number of layers and neurons per layer affects model capacity.

Batch Size

Controls gradient noise and memory usage. Smaller batches can escape local minima.

Interpreting Test Results

Analyzing Model Strengths

- Classes with high precision and recall
- Consistent performance across data subsets
- Robust to minor input variations

Identifying Weaknesses

- Classes with poor performance metrics
- Sensitivity to specific transformations
- Performance gaps between demographic groups

Always evaluate models on metrics that align with your application's goals.

Future Scope: Explainable AI

Model Transparency

Making deep learning models more interpretable to humans through visualization techniques and explanation methods.

Ethical AI

Understanding model decisions helps identify and mitigate bias, ensuring fair and ethical AI applications.

User Trust

Explainable models build trust with users and stakeholders in sensitive domains like healthcare and finance.

Next week, we'll explore deploying models to production environments!