Categorical Cross-Entropy

Loss Function for Multiclass Classification

Formula:

$$L = -\sum_k y_k \log(\hat{y}_k)$$

One-hot Encoding:

```
y = [0, 0, 1, 0, ...]
```

for true class

Combined with Softmax:

Softmax + CCE

Differentiable end-to-end

Σ Generalizes binary cross-entropy to **K classes**

⚠ Penalizes deviation from true class distribution

Gradient:

$$\hat{y}_k - y_k$$

(elegant and simple!)

Binary → 2 classes
Categorical → K classes