# Multi-class Metrics

## Macro-Average

$$\text{avg}(\text{metric}_1, \text{metric}_2, \ldots)$$

Calculate metric for each class, then average. Equal class weight.

## Micro-Average

$$\text{metric}(\Sigma\ TP, \Sigma\ FP, \Sigma\ FN)$$

Aggregate all classes first, then calculate. Favors majority.

## Weighted-Average

$$\Sigma(\text{weight} \times \text{metric})$$

Weight by class support. Balances macro & micro.

### Example: F1 Score Calculation

| | |
|---|---|
| Class A (100 samples) | F1 = 0.90 |
| Class B (50 samples) | F1 = 0.70 |
| Class C (50 samples) | F1 = 0.60 |

**Macro-Average**
(0.90 + 0.70 + 0.60) / 3 → **0.73**

**Micro-Average**
Aggregate TP/FP/FN → **0.82**

**Weighted-Average**
(100×0.90+50×0.70+50×0.60)/200 → **0.78**

## 🔑 Key Insights

**Macro**
Treats all classes equally, good for balanced view

**Micro**
Similar to accuracy, biased to majority class

**Weighted**
Balanced approach considering class sizes

### Confusion Matrix (5 Classes)

| | C0 | C1 | C2 | C3 | C4 |
|---|---|---|---|---|---|
| **C0** | **45** | 3 | 2 | 0 | 0 |
| **C1** | 2 | **28** | 5 | 0 | 0 |
| **C2** | 1 | 4 | **18** | 2 | 0 |
| **C3** | 0 | 0 | 3 | **12** | 0 |
| **C4** | 0 | 0 | 1 | 1 | **8** |

Rows: Actual | Columns: Predicted | Total: 135 samples

### Per-Class Metrics

| Class | Support | Precision | Recall | F1 |
|---|---|---|---|---|
| C0 | 50 | 0.938 | 0.900 | **0.918** |
| C1 | 35 | 0.800 | 0.800 | **0.800** |
| C2 | 25 | 0.621 | 0.720 | **0.667** |
| C3 | 15 | 0.800 | 0.800 | **0.800** |
| C4 | 10 | 1.000 | 0.800 | **0.889** |

**Example Calculation for C0:**

- TP = 45, FP = 3, FN = 5
- Precision = 45/(45+3) = 0.938
- Recall = 45/(45+5) = 0.900
- F1 = 2×(0.938×0.900)/(0.938+0.900) = 0.918

## 📊 Final Averaged Metrics

### Macro-Average F1

**0.815**

(0.918 + 0.800 + 0.667
+ 0.800 + 0.889) / 5

### Micro-Average F1

**0.822**

ΣTP = 111, ΣFP = 24, ΣFN = 24
F1 = 2×111 / (2×111+24+24)

### Weighted-Average F1

**0.835**

(50×0.918 + 35×0.800 + 25×0.667
+ 15×0.800 + 10×0.889) / 135

💡 **Interpretation:** Weighted (0.835) > Micro (0.822) > Macro (0.815) indicates the model performs better on majority classes (C0, C1) than on minority classes