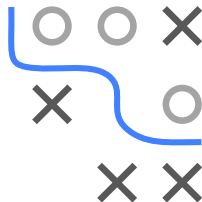


Introduction to Machine Learning

Evaluation

Multi-Class AUC



AUC(pos|neg) = AUC(1|3)

| | Y | $\hat{\pi}_1$ | $\hat{\pi}_2$ | $\hat{\pi}_3$ |
|-----|---|---------------|---------------|---------------|
| pos | 1 | 0.7 | 0.2 | 0.1 |
| pos | 1 | 0.5 | 0.3 | 0.2 |
| neg | 2 | 0.3 | 0.5 | 0.2 |
| neg | 2 | 0.4 | 0.5 | 0.1 |
| neg | 3 | 0.6 | 0.1 | 0.3 |
| neg | 3 | 0.1 | 0.1 | 0.8 |

AUC(pos|neg) = AUC(3|1)

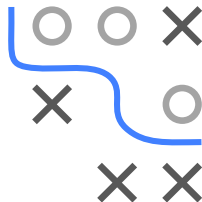
| | Y | $\hat{\pi}_1$ | $\hat{\pi}_2$ | $\hat{\pi}_3$ |
|-----|---|---------------|---------------|---------------|
| neg | 1 | 0.7 | 0.2 | 0.1 |
| neg | 1 | 0.5 | 0.3 | 0.2 |
| pos | 2 | 0.3 | 0.5 | 0.2 |
| pos | 2 | 0.4 | 0.5 | 0.1 |
| pos | 3 | 0.6 | 0.1 | 0.3 |
| pos | 3 | 0.1 | 0.1 | 0.8 |

Learning goals

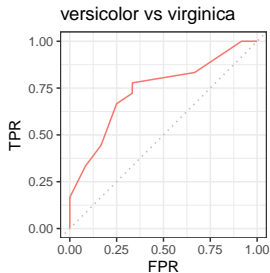
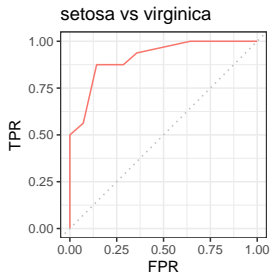
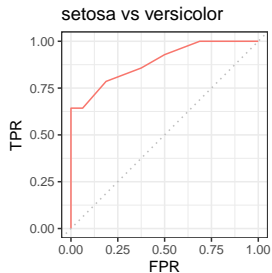
- Understand that generalizing AUC to multi-class is not trivial
- Learn how multi-class AUC can be derived

MULTI-CLASS AUC

- AUC and other ROC metrics for binary classification
- Different ways to estimate **multi-class AUC**
- Often based on aggregated binary AUCs:
e.g. 1-vs-1 or 1-vs-rest

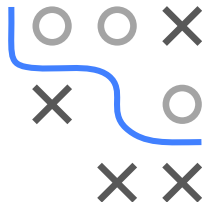


Example: 1-vs-1 on iris



MULTI-CLASS AUC

- For binary classes: always $AUC(1|0) = AUC(0|1)$
- For multi-class usually: $AUC(k | \ell) \neq AUC(\ell | k)$
- **Example** with $g = 3$ where $AUC(1|3) \neq AUC(3|1)$:
 - $AUC(3|1) = 1$ (RHS) as before
 - $AUC(1|3) \neq 1$ (LHS)



AUC(pos|neg) = AUC(1|3)

| | Y | $\hat{\pi}_1$ | $\hat{\pi}_2$ | $\hat{\pi}_3$ |
|----------------|--------------|----------------|----------------|----------------|
| pos | 1 | 0.7 | 0.2 | 0.1 |
| pos | 1 | 0.5 | 0.3 | 0.2 |
| pos | 2 | 0.3 | 0.5 | 0.2 |
| pos | 2 | 0.4 | 0.5 | 0.1 |
| neg | 3 | 0.6 | 0.1 | 0.3 |
| neg | 3 | 0.1 | 0.1 | 0.8 |

AUC(pos|neg) = AUC(3|1)

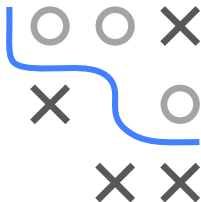
| | Y | $\hat{\pi}_1$ | $\hat{\pi}_2$ | $\hat{\pi}_3$ |
|----------------|--------------|----------------|----------------|----------------|
| neg | 1 | 0.7 | 0.2 | 0.1 |
| neg | 1 | 0.5 | 0.3 | 0.2 |
| neg | 2 | 0.3 | 0.5 | 0.2 |
| neg | 2 | 0.4 | 0.5 | 0.1 |
| pos | 3 | 0.6 | 0.1 | 0.3 |
| pos | 3 | 0.1 | 0.1 | 0.8 |

MULTI-CLASS AUC

Hand and Till (2001) proposed to avg AUC via **1-vs-1**:

- For all class pairs, compute $AUC(k | \ell)$.

$$AUC_{MC} = \frac{1}{g(g-1)} \sum_{k \neq \ell} AUC(k|\ell) \in [0, 1].$$



Comments:

- Other defs use **1-vs-rest** and need to avg only g AUC values
- 1-vs-rest creates imbal classes even if orig classes are balanced
- Imbalanced classes can be considered by weighting individual AUC values with class priors [Ferri et al. (2003)]