

Binary classification

Tim Lawson

October 20, 2023

Binary relations

If A and B are sets:

- ▶ The Cartesian product $A \times B$ is the set of pairs $\{(x, y) \mid x \in A, y \in B\}$.
- ▶ A binary relation is a set of pairs $R \subseteq A \times B$.
- ▶ If $A = B$, then the relation is “over A ”.
- ▶ Instead of $(x, y) \in R$, we also write xRy .

Binary relations

- ▶ **Reflexive** if $xRx \forall x \in A$

For all x in A , (x, x) is in R .

- ▶ **Symmetric** if $xRy \implies yRx \forall x, y \in A$

For all x, y in A , if (x, y) is in R , then (y, x) is in R .

- ▶ **Antisymmetric** if $xRy \wedge yRx \implies x = y \forall x, y \in A$

For all x, y in A , if (x, y) and (y, x) are in R , then $x = y$.

- ▶ **Transitive** if $xRy \wedge yRz \implies xRz \forall x, y, z \in A$

For all x, y, z in A , if (x, y) and (y, z) are in R , then (x, z) is in R .

- ▶ **Total** if $xRy \vee yRx \forall x, y \in A$

For all x, y in A , (x, y) or (y, x) is in R .

If a binary relation is total, then it is also reflexive.

Partial orders

- ▶ A **partial order** is a binary relation that is reflexive, antisymmetric and transitive.
- ▶ For instance, the **subset** relation \subseteq on sets is a partial order:

✓ Reflexive: $A \subseteq A \ \forall \ A$

A is a subset of itself.

✓ Antisymmetric: $A \subseteq B \wedge B \subseteq A \implies A = B$

If A is a subset of B and B is a subset of A , then $A = B$.

✓ Transitive: $A \subseteq B \wedge B \subseteq C \implies A \subseteq C$

If A is a subset of B and B is a subset of C , then A is a subset of C .

Total orders

- ▶ A **total order** is a binary relation that is total, antisymmetric and transitive.
- ▶ For instance, the \leq relation on real numbers is a total order:
 - ✓ Total: $x \leq y \vee y \leq x \quad \forall x, y \in \mathbb{R}$
 - ✓ Antisymmetric: $x \leq y \wedge y \leq x \implies x = y$
 - ✓ Transitive: $x \leq y \wedge y \leq z \implies x \leq z$

Equivalence relations

- ▶ An **equivalence relation** is a binary relation \equiv that is reflexive, symmetric and transitive.
- ▶ For instance, the relation 'contains the same number of elements as' on sets, i.e., $|A| = |B|$, is an equivalence relation:
 - ✓ Reflexive: $|A| = |A| \forall A$
 - ✓ Antisymmetric: $|A| = |B| \wedge |B| = |A| \implies |A| = |B|$
 - ✓ Transitive: $|A| = |B| \wedge |B| = |C| \implies |A| = |C|$

Measures

- ▶ “To achieve good accuracy, a classifier should concentrate on the *majority class*, particularly if the class distribution is highly unbalanced”¹
- ▶ “If the minority class is the class of interest and very small, accuracy and performance on the majority class are not the right quantities to optimise”²

¹Flach 2012, p.56.

²Flach 2012, p.57.

Coverage plots

- ▶ “If one classifier outperforms another classifier on all classes, the first one is said to dominate the second”³
 - ▶ More true positives and fewer false positives
 - ▶ Above and to the left
- ▶ “Which one we prefer depends on whether we put more emphasis on the positives or on the negatives”⁴
- ▶ Demonstration

³Flach 2012, p.59.

⁴Flach 2012, p.59.

Bibliography



Flach, Peter (2012). *Machine Learning: The Art and Science of Algorithms That Make Sense of Data*. 1st ed. Cambridge University Press.