Multiclass classification with binary classifiers

Suppose you have $k$ classes but want to train binary classifiers.

Two main methods:

- **One vs. All** (also called “One vs. Rest”)
- **All-pairs** (also called “One vs. One”)
One vs. All method

Training:

Train $k$ classifiers.

Example: For handwritten digits task:

- ‘0’ vs. not ‘0’
- ‘1’ vs. ‘not ‘1’
- etc.

Testing:

Given instance $x$:

- Run each classifier on $x$.
- If class is positive, record the decision value: $d = w \cdot x$
- Return the class with highest decision value (breaking ties at random)
All-Pairs method

Training:

Train separate perceptron for each pair of classes. For \( k \) classes, \((k (k - 1) / 2)\) perceptrons.

- ‘0’ vs. ‘1’
- ‘1’ vs. ‘2’
- ‘0’ vs. ‘2’
- ‘1’ vs. ‘3’
- ‘0’ vs. ‘3’
- ‘1’ vs. ‘4’
- etc.

Testing:

Given instance \( x \):
- Run each perceptron on \( x \). Collect “votes” for each class.
- Return class that received most votes (breaking ties at random)
Example of All-Pairs method

Suppose you are classifying ‘A’ vs. ‘B’ vs. ‘C’ vs. ‘D’

A-B classifier says ‘A’

A-C classifier says ‘C’

A-D classifier says ‘A’

B-C classifier says ‘B’

B-D classifier says ‘D’

C-D classifier says ‘D’

What is the class?