Binary classification

$$C = \begin{pmatrix} t_n & f_p \\ f_n & t_p \end{pmatrix} \begin{matrix} 1 \\ 2 \end{matrix} \quad \vec{t}$$
$$\qquad\quad 1 \quad 2 \qquad\quad \vec{y}$$

* Two classes
  - negative → class 1
  + positive → class 2

∴ negatives = $n = \#(\vec{t}=1)$

∴ positives = $p = \#(\vec{t}=2)$

negative rate (frequency) $p_n = \frac{\#(\vec{t}=1)}{N} = \frac{n}{N}$

positive rate (frequency), $p_p = \frac{\#(\vec{t}=2)}{N} = \frac{p}{N}$

true negatives: $t_n = \#(\vec{y}=1, \vec{t}=1)$

false negatives: $f_n = \#(\vec{y}=1, \vec{t}=2)$

true positives: $t_p = \#(\vec{y}=2, \vec{t}=2)$

false positives: $f_p = \#(\vec{y}=2, \vec{t}=1)$

true negative rate, $t_{nr} = \frac{\#(\vec{y}=1, \vec{t}=1)}{\#(\vec{t}=1)} = p(\vec{y}=1 \mid \vec{t}=1)$

false negative rate, $f_{nr} = p(\vec{y}=1 \mid \vec{t}=2) = \frac{\#(\vec{y}=1, \vec{t}=2)}{\#(\vec{t}=2)}$

true positive rate, $t_{pr} = p(\vec{y}=2 \mid \vec{t}=2) = \frac{\#(\vec{y}=2, \vec{t}=2)}{\#(\vec{t}=2)}$

false positive rate, $f_{pr} = p(\vec{y}=2 \mid \vec{t}=1) = \frac{\#(\vec{y}=2, \vec{t}=1)}{\#(\vec{t}=1)}$

$$t_{pr} + f_{nr} = 1$$
$$t_{nr} + f_{pr} = 1$$

classification error, $E = p(\vec{y} \neq \vec{t})$

$$= p(\vec{y}=1, \vec{t}=2) + p(\vec{y}=2, \vec{t}=1)$$

$$E = f_{nr} \cdot p_p + f_{pr} \cdot p_n$$

∴ Accuracy

$$A = 1 - E$$

Sensitivity / recall $\Rightarrow$ tpr

Specificity / Selectivity $\Rightarrow$ tnr

Precision $= P(\vec{t}=1 \mid \vec{y}=1)$

$$= \frac{t_p}{f_p + t_p}$$

$$\text{Precision} = \frac{tpr}{fpr \cdot \frac{P_n}{P_p} + tpr}$$

CNN dimension

$$H^{(l)} = 1 + \frac{H^{(l-1)} - f_y}{\Delta y}$$

$$W^{(l)} = 1 + \frac{W^{(l-1)} - f_x}{\Delta x}$$

Gradient descent

$$\vec{x}^{(t+1)} = \vec{x}^{(t)} - \varepsilon \vec{\nabla} h \Big|_{\vec{x}^{(t)}}$$

DNN

$$A = XW + \vec{b}^T$$

In general,

$$A^{(i)} = A^{(i-1)} \cdot W^{(i)} + \vec{b}^{(i)T}$$

- np.random.randint
- support points, np.linspace()
- uniform distribution, np.random.uniform()