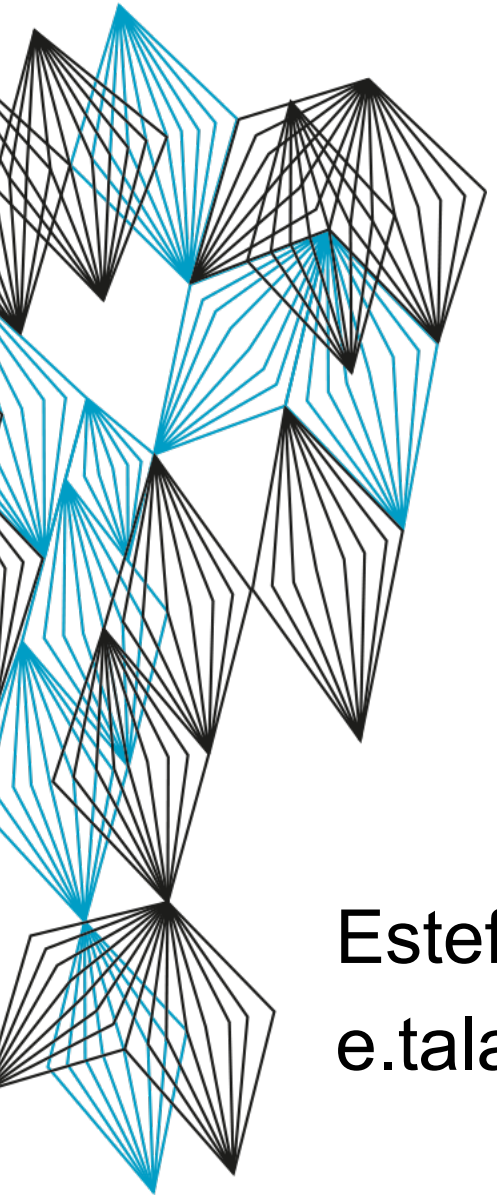


ARTIFICIAL INTELLIGENCE & CYBER SECURITY

INTRO TO MACHINE LEARNING VALIDATION

Estefanía Talavera Martínez
e.talaveramartinez@utwente.nl

What error measure (or utility function)?

Different problems require different types of ***Error Functions***:

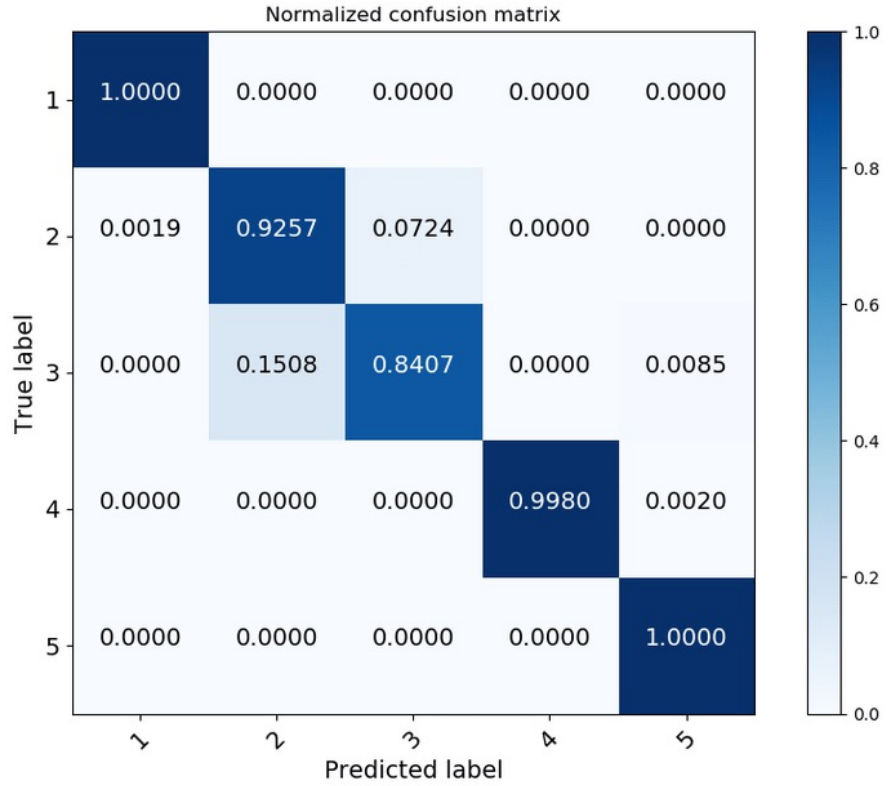
Supervised learning
Classification Some form of counting misclassified datapoints
Regression Average distance between predicted and target output

Unsupervised learning Within-class and between-class distances

Errors for Supervised Learning - Classification

The primary source for performance estimation is the confusion matrix TP, TN, FP, FN

Predicted class \ True class	Positive	Negative
Positive	True Positive count (TP)	False Positive count (FP)
Negative	False Negative count (FN)	True Negative count (TN)



$$Acc = \frac{TP+TN}{TP+TN+FP+FN} \qquad Precision = \frac{TP}{TP+FP} \qquad F_{score} = 2\frac{Precision*Recall}{Precision+Recall}$$

$$Recall/Sensitivity/TruePositiveRate = \frac{TP}{TP+FN} \qquad Specificity/TrueNegativeRate = \frac{TN}{TN+FP}$$

Errors for Supervised Learning - Regression

Assess the difference between predicted output and target output

Sum of squared errors (SSE)
Which we want to minimise

$$E = \frac{1}{2}\sum_{n=1}^{N}(y(\mathrm{x}_n) - t_n)^2$$

Probability of the predicted outputs given the target outputs
Which we want to maximise

Errors for Unsupervised Learning - Clustering

Hard problem
In most applications, expert judgements are still the key.

Internal measures - Cohesion vs Separation

Cohesion: how closely related are samples in a cluster

Within sum squared errors (WSS)

$$WSS = \sum_i \sum_{x \in C_i} (x - m_i)^2$$

Separation: how well separated is one cluster from other clusters

Between cluster sum of squares (BSS)

$$BSS = \sum_i |C_i| (m - m_i)^2$$

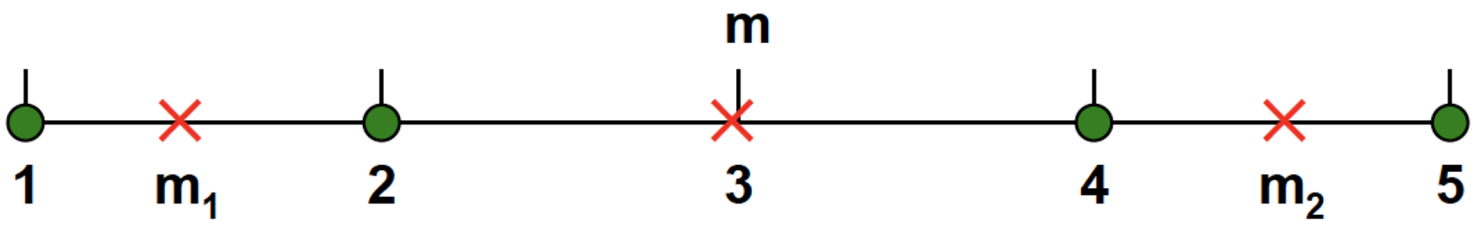
$|C_i|$ = size of the cluster

Example

$$WSS = \sum_i \sum_{x \in C_i} (x - m_i)^2$$

$$BSS = \sum_i |C_i| (m - m_i)^2$$

- BSS + WSS = constant



K=1 cluster:

$$WSS = (1-3)^2 + (2-3)^2 + (4-3)^2 + (5-3)^2 = 10$$

$$BSS = 4 \times (3-3)^2 = 0$$

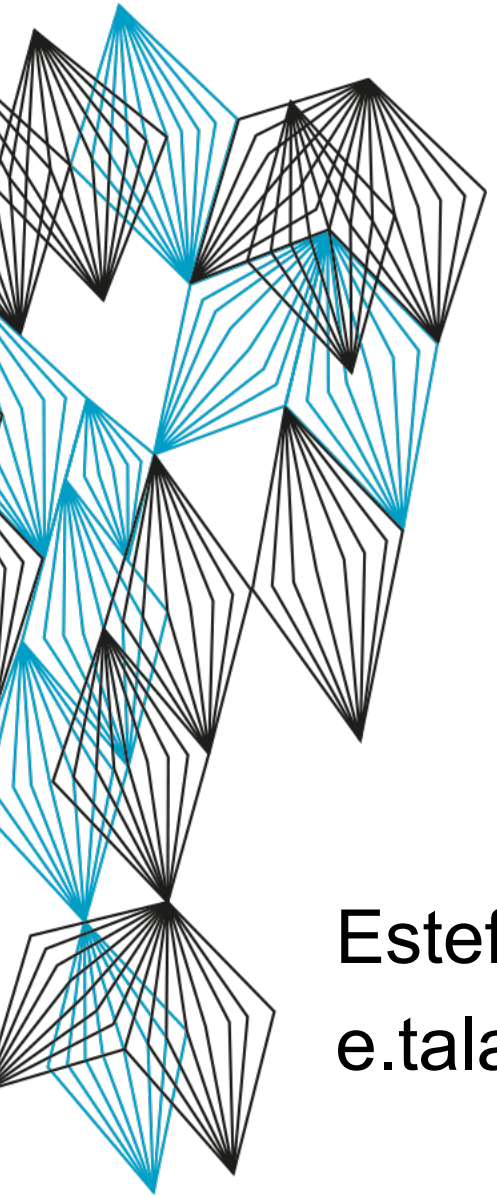
$$Total = 10 + 0 = 10$$

K=2 clusters:

$$WSS = (1-1.5)^2 + (2-1.5)^2 + (4-4.5)^2 + (5-4.5)^2 = 1$$

$$BSS = 2 \times (3-1.5)^2 + 2 \times (4.5-3)^2 = 9$$

$$Total = 1 + 9 = 10$$



ARTIFICIAL INTELLIGENCE & CYBER SECURITY

INTRO TO MACHINE LEARNING VALIDATION

Estefanía Talavera Martínez
e.talaveramartinez@utwente.nl