

Metrics for Supervised Learning

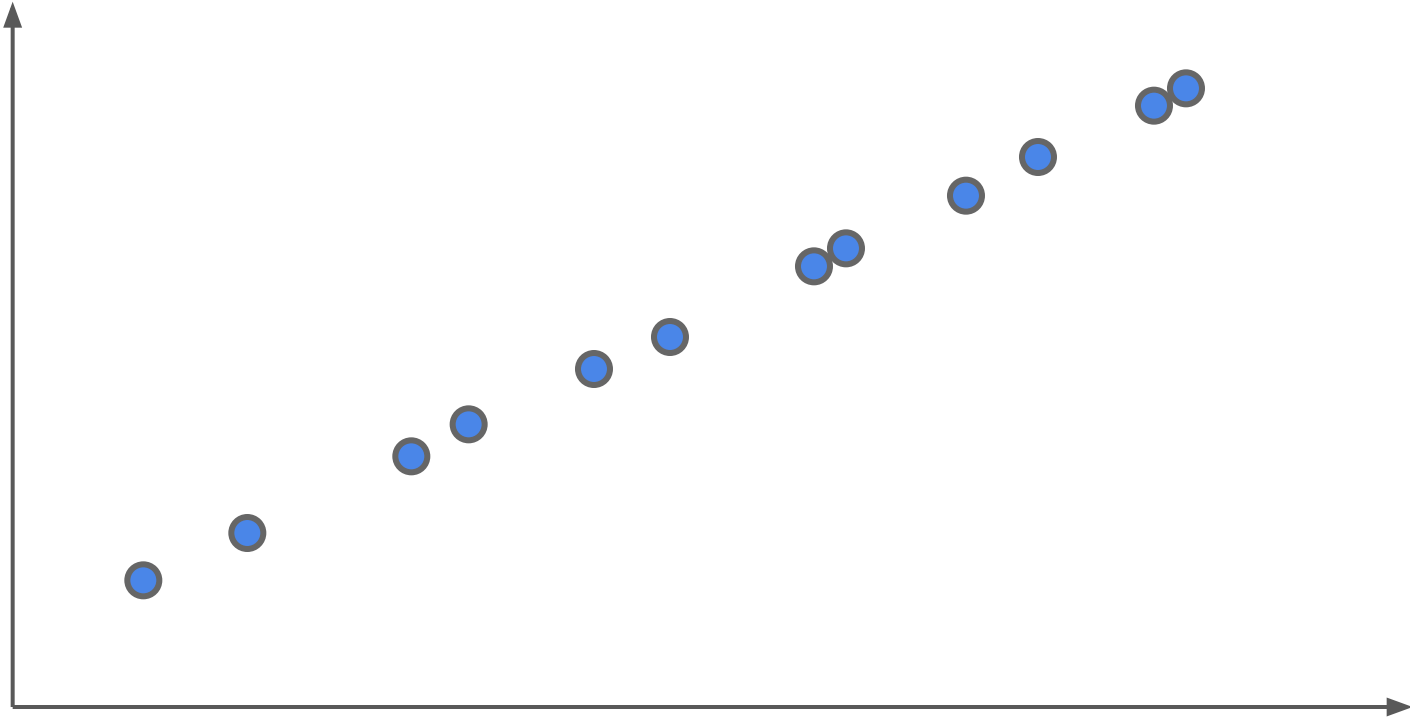
MACHINE LEARNING

Pakarat Musikawan

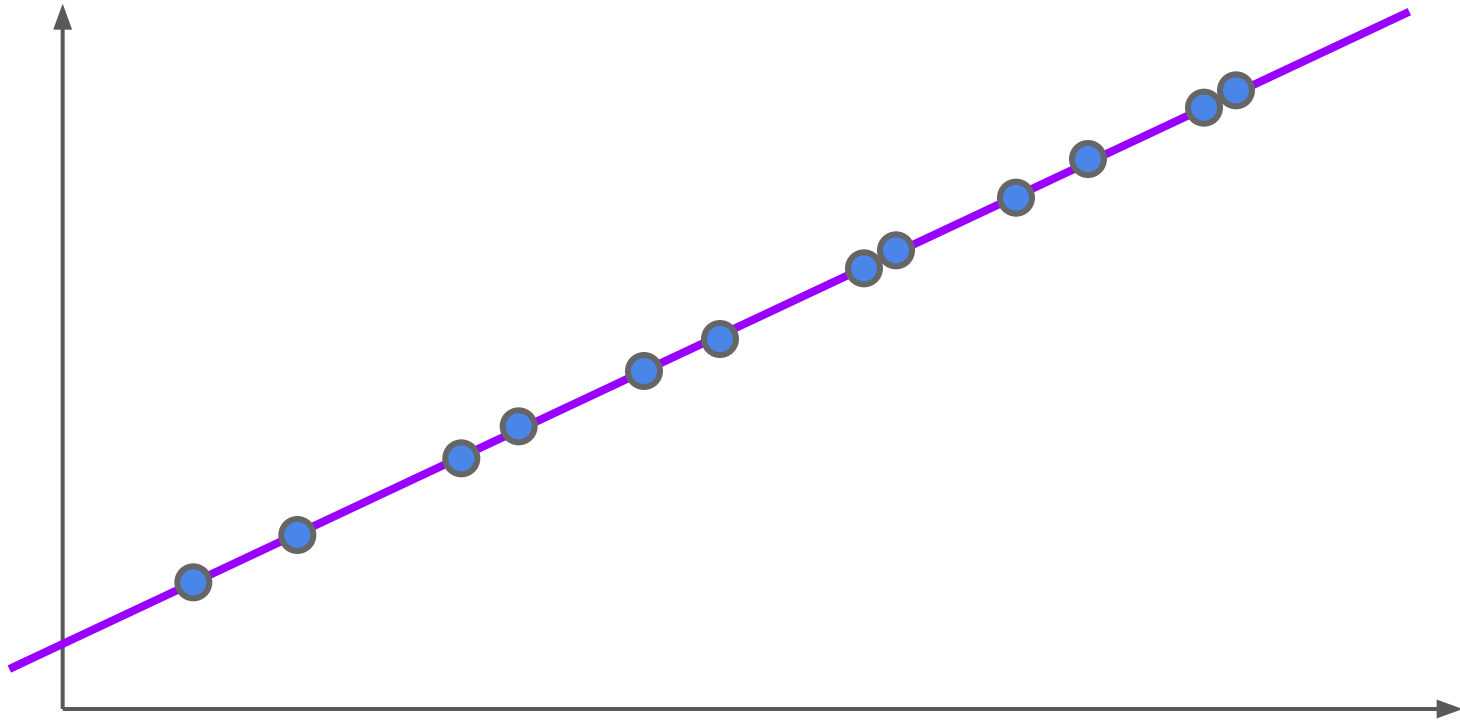
Metrics

- Regression/Estimation
 - Mean squared error (MSE)
 - Root mean squared error (RMSE)
 - Mean absolute error (MAE)
 - Mean absolute percentage error (MAPE)
 - R-squared (R^2 or the coefficient of determination)
- Classification
 - Accuracy
 - Precision
 - Recall
 - F1-score
 - Receiver operating characteristic (ROC)
 - Area under the ROC Curve (AUC)

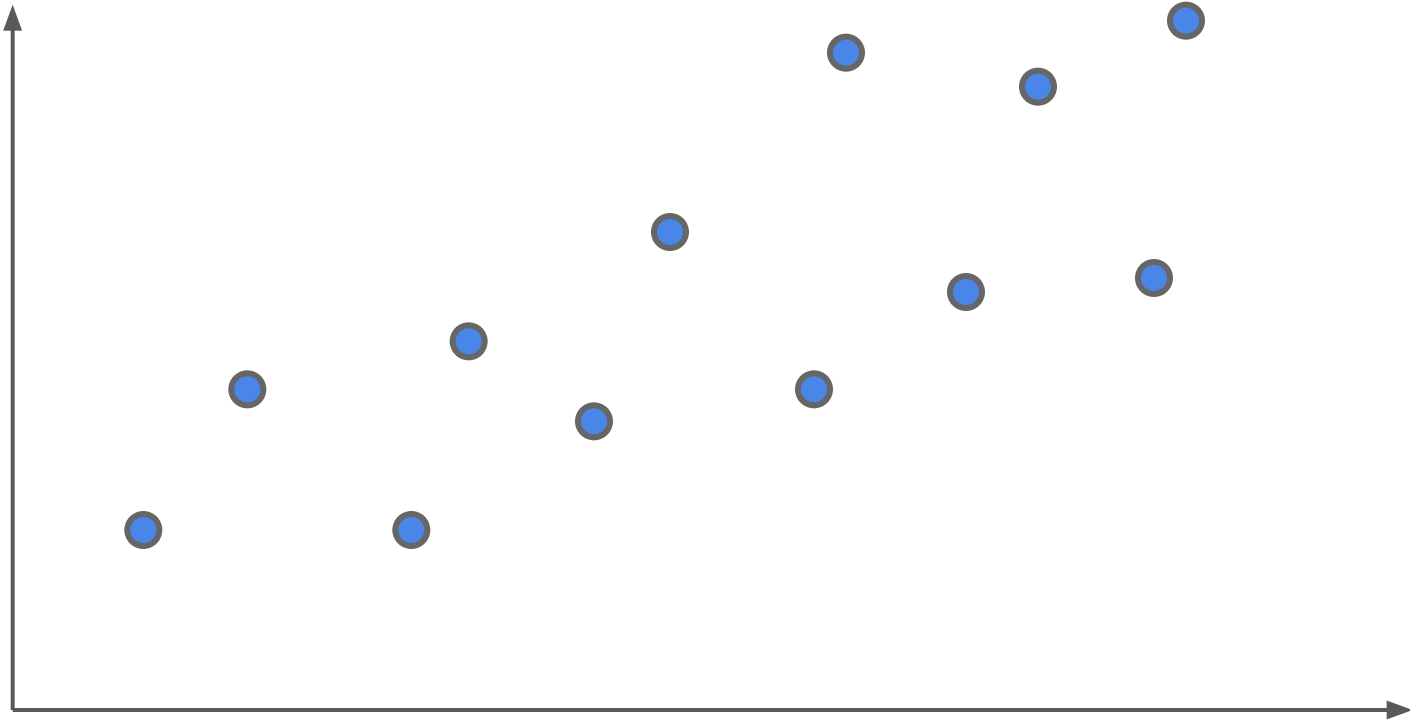
Model Evaluation in Regression/Estimation



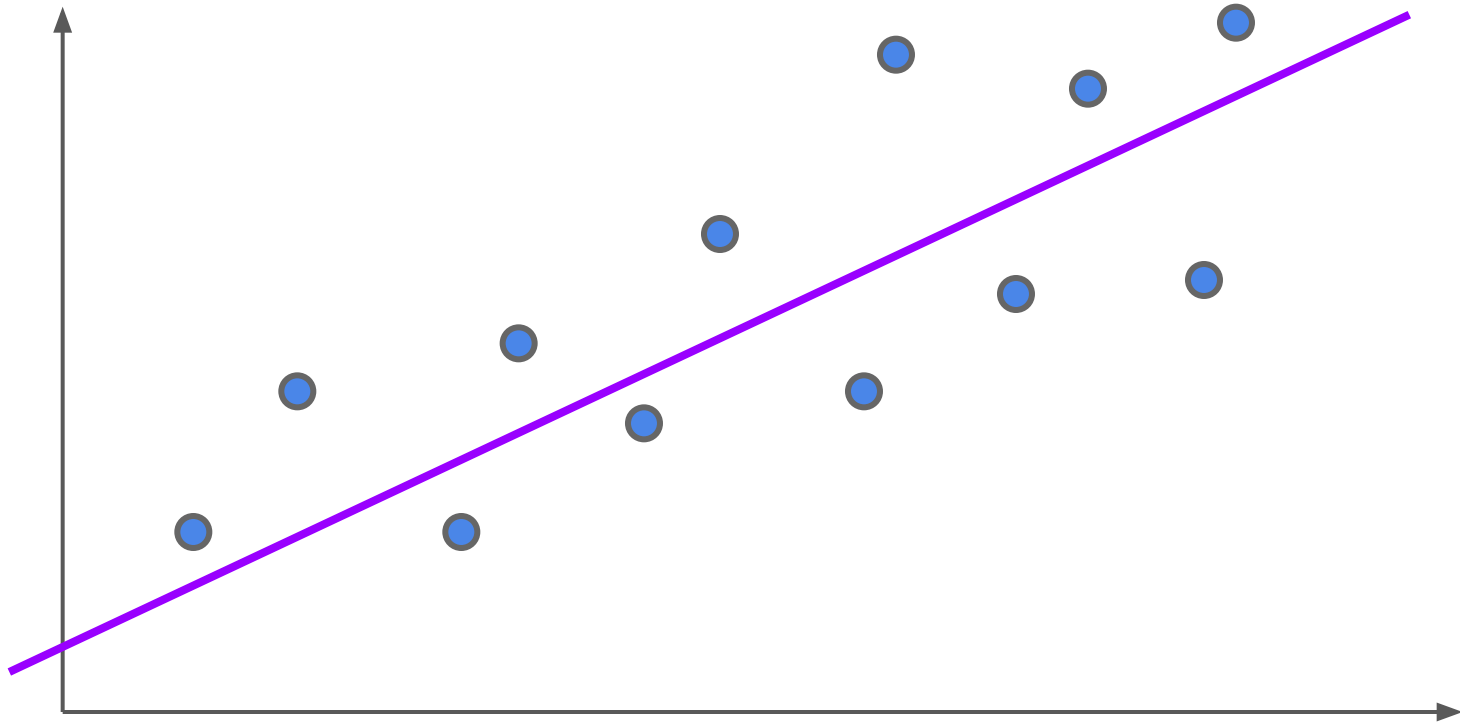
Model Evaluation in Regression/Estimation



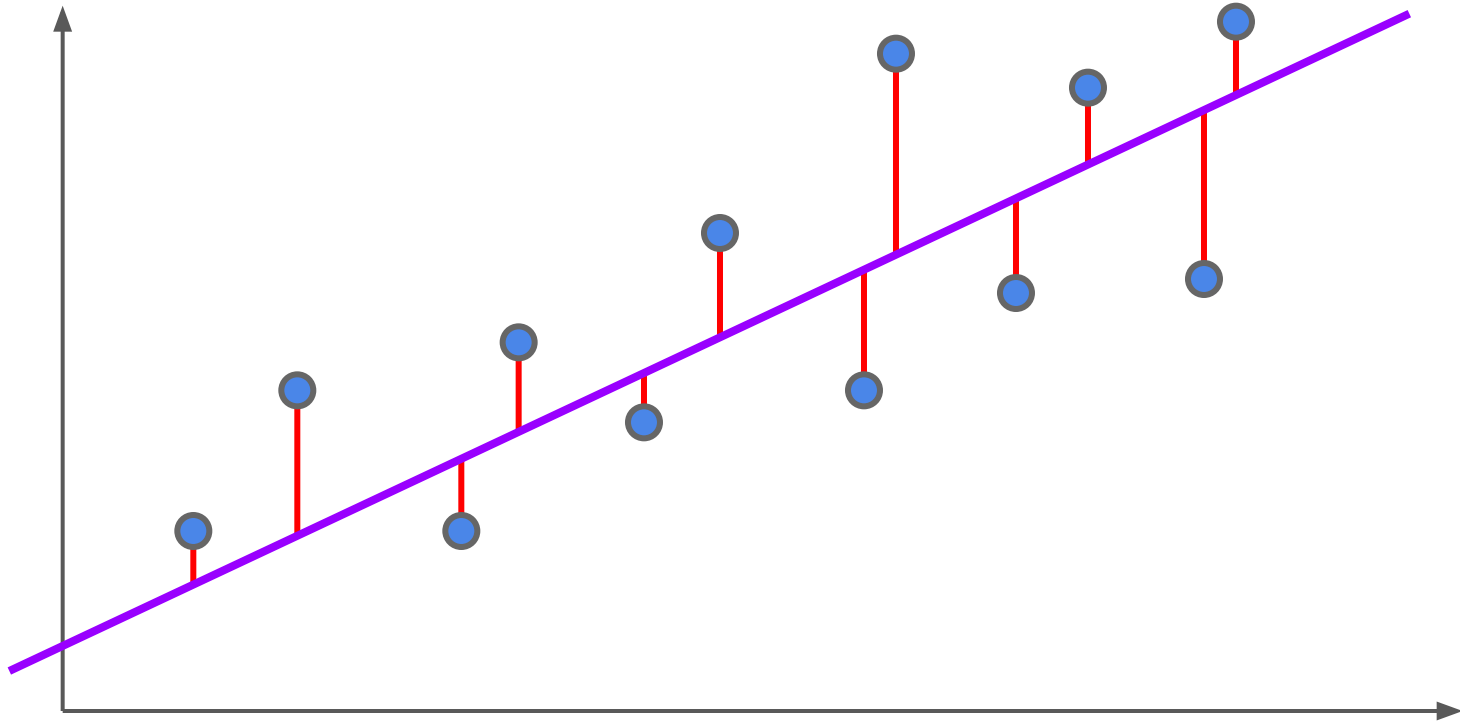
Model Evaluation in Regression/Estimation



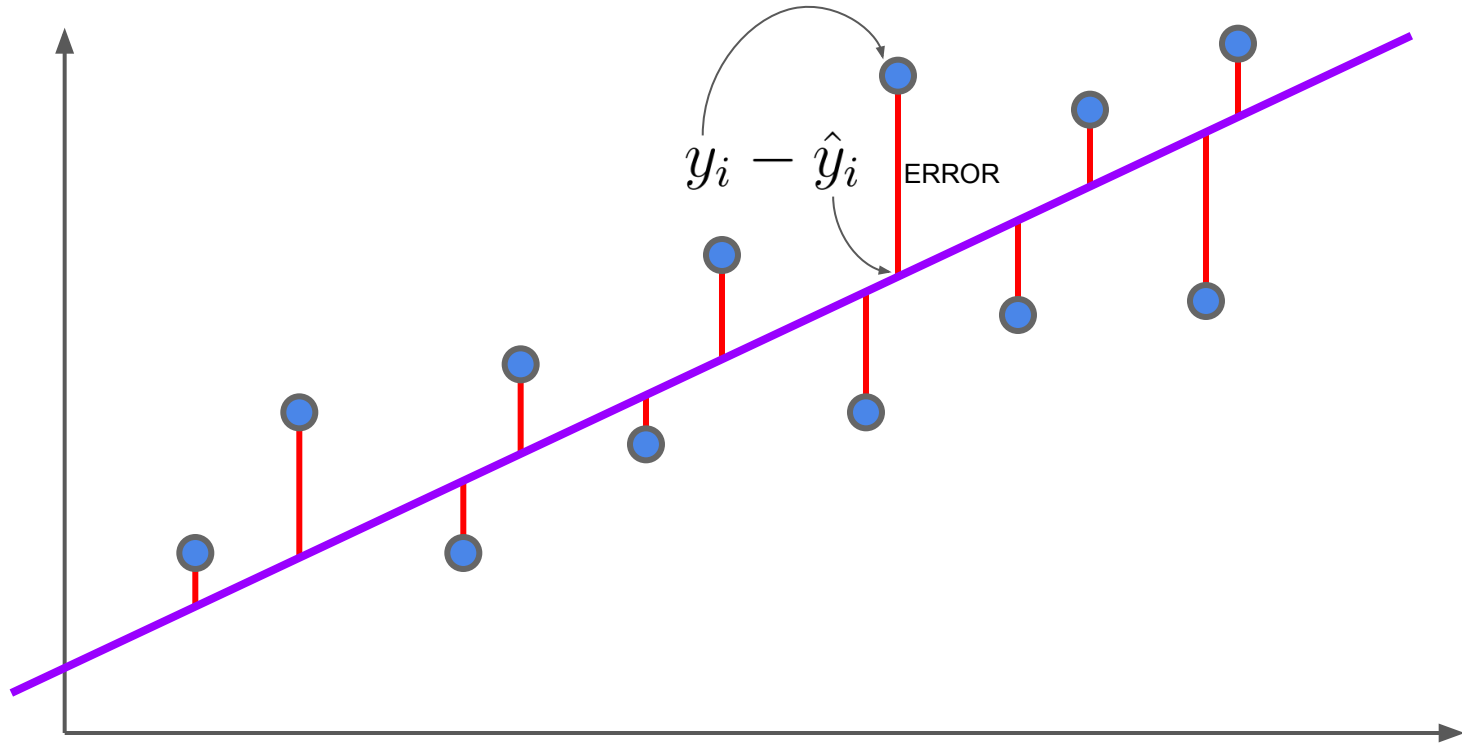
Model Evaluation in Regression/Estimation



Model Evaluation in Regression/Estimation



Model Evaluation in Regression/Estimation



Model Evaluation in Regression/Estimation

$\hat{\mathbf{y}}$ = Predicted values

\mathbf{y} = True values

$\hat{\mathbf{y}} = [\hat{y}_1, \hat{y}_2, \dots, \hat{y}_N]$

$\mathbf{y} = [y_1, y_2, \dots, y_N]$

$$\text{MSE} = \frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2$$

$$\text{MAE} = \frac{1}{N} \sum_{i=1}^N |y_i - \hat{y}_i|$$

$$\text{MAPE} = \frac{1}{N} \sum_{i=1}^N \frac{|y_i - \hat{y}_i|}{y_i}$$

$$R^2 = 1 - \frac{\sum_{i=1}^N (y_i - \hat{y}_i)^2}{\sum_{i=1}^N (y_i - \bar{y})^2}$$

$$\hat{\mathbf{y}} = [0.0, 0.2, 0.4]$$

$$\mathbf{y} = [0.1, 0.3, 0.5]$$

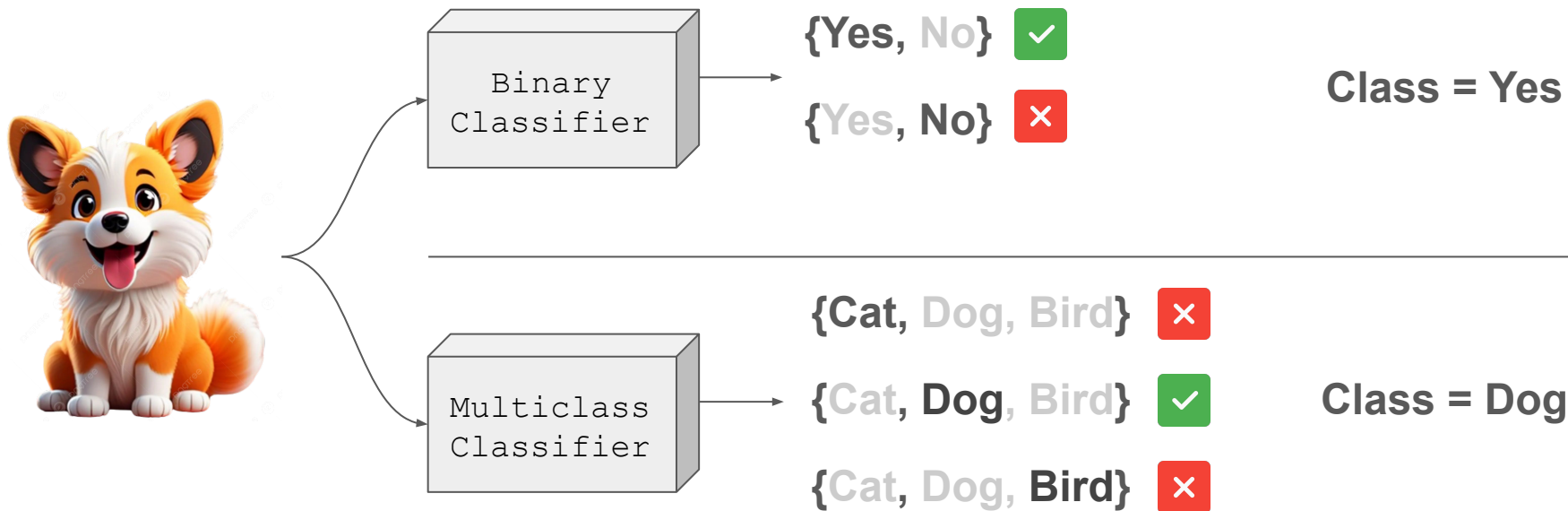
$$\text{MSE} = \frac{(0.1 - 0.0)^2 + (0.3 - 0.2)^2 + (0.5 - 0.4)^2}{3} = 0.01$$

$$\text{MAE} = \frac{|0.1 - 0.0| + |0.3 - 0.2| + |0.5 - 0.4|}{3} = 0.1$$

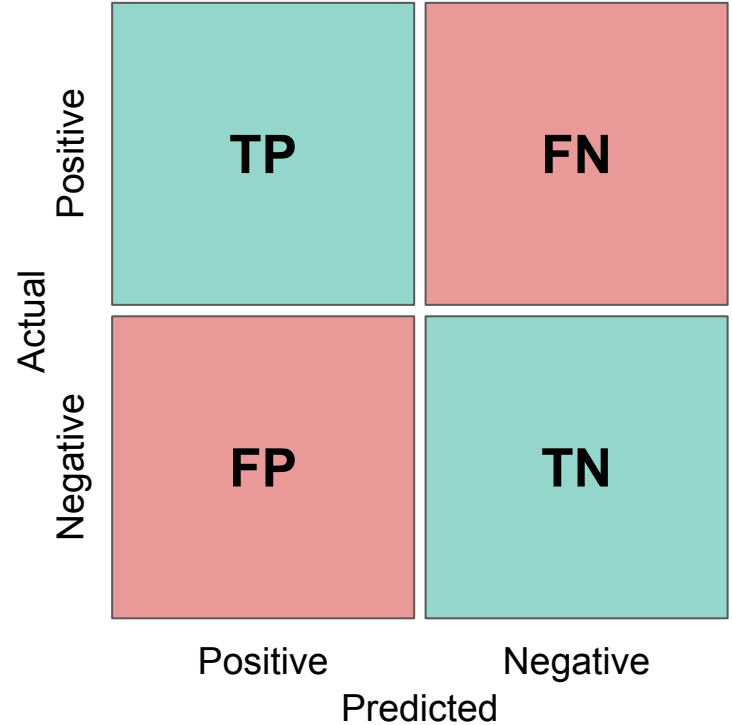
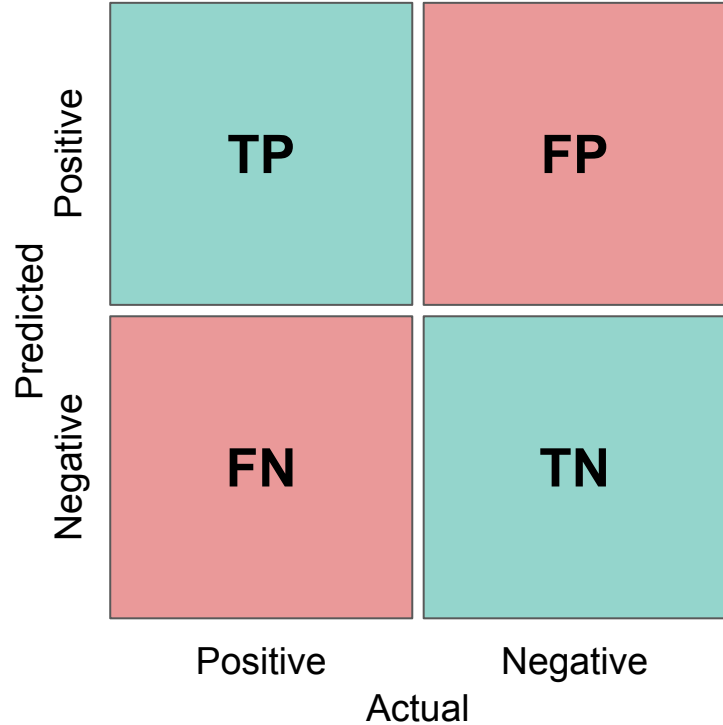
$$\text{MAPE} = \frac{\frac{|0.1-0.0|}{0.1} + \frac{|0.3-0.2|}{0.3} + \frac{|0.5-0.4|}{0.5}}{3} = 0.511$$

$$R^2 = 1 - \frac{(0.1 - 0.0)^2 + (0.3 - 0.2)^2 + (0.5 - 0.4)^2}{(0.1 - 0.3)^2 + (0.3 - 0.3)^2 + (0.5 - 0.3)^2} = 0.75$$

Model Evaluation in Classification



Model Evaluation in Binary Classification



Model Evaluation in Binary Classification

Predicted	Positive	TP	FP
	Negative	FN	TN
		Positive	Negative
		Actual	

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FP} + \text{TN} + \text{FN}}$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}$$

$$\text{F1-score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

Model Evaluation in Binary Classification

Predicted	Positive	4	1
	Negative	2	3
		Positive	Negative
		Actual	

1: Positive, 0:Negative

y = True class

\hat{y} = Predicted class

$y = [y_1, y_2, \dots, y_N]$

$\hat{y} = [\hat{y}_1, \hat{y}_2, \dots, \hat{y}_N]$

$y = [1, 1, 0, 1, 0, 1, 0, 1, 0, 1]$

$\hat{y} = [1, 0, 0, 1, 0, 1, 0, 1, 1, 0]$

Model Evaluation in Binary Classification

Predicted	Positive	Negative
	Actual	Actual
Positive	189	26
Negative	51	234

หากเราทำการสร้างตัวแบบสำหรับการจำแนกข้อมูลของการตัดสินใจซื้อรถของลูกค้า

โดยข้อมูลดังกล่าวมีจำนวนทั้งหมด 500 ข้อมูล และสามารถแบ่งออกเป็น 2 ประเภท ได้แก่ข้อมูล “ซื้อ” มีจำนวน 240 ข้อมูล และข้อมูล “ไม่ซื้อ” มีจำนวน 260 ข้อมูล

ผลปรากฏว่าตัวแบบที่เราสร้างขึ้นสามารถทำการจำแนกข้อมูล “ซื้อ” ได้ถูกต้องเป็นจำนวน 189 ข้อมูล และสามารถทำการจำแนกข้อมูล “ไม่ซื้อ” ได้ถูกต้องเป็นจำนวน 234 ข้อมูล จึงทำการสร้าง Confusion Matrix สำหรับผลลัพธ์ดังกล่าว

- “ซื้อ” เป็น Positive
- “ไม่ซื้อ” เป็น Negative

Model Evaluation in Binary Classification

Predicted	Positive	Negative
	189	26
Actual		
Positive	51	234
Negative		

$$\begin{aligned}\text{Accuracy} &= \frac{TP + TN}{TP + TN + FP + FN} \\ &= \frac{189 + 234}{189 + 234 + 26 + 51} \\ &= 0.846\end{aligned}$$

$$\begin{aligned}&= \frac{\sum_{i=1}^N \mathbb{I}(\hat{y}_i = y_i)}{N} \\ &= \frac{423}{500} = 0.846\end{aligned}$$

Model Evaluation in Binary Classification

Predicted	Positive	Negative
	Actual	Actual
Positive	189	26
Negative	51	234

$$\begin{aligned}\text{Precision} &= \frac{TP}{TP + FP} \\ &= \frac{189}{189 + 26} \\ &= 0.879\end{aligned}$$

Precision:
ความน่าจะเป็นที่โมเดล
ทำนายข้อมูล
Positive ถูกต้องจาก
การทำนาย Positive
ทั้งหมด

$$\begin{aligned}\text{Recall} &= \frac{TP}{TP + FN} \\ &= \frac{189}{189 + 51} \\ &= 0.788\end{aligned}$$

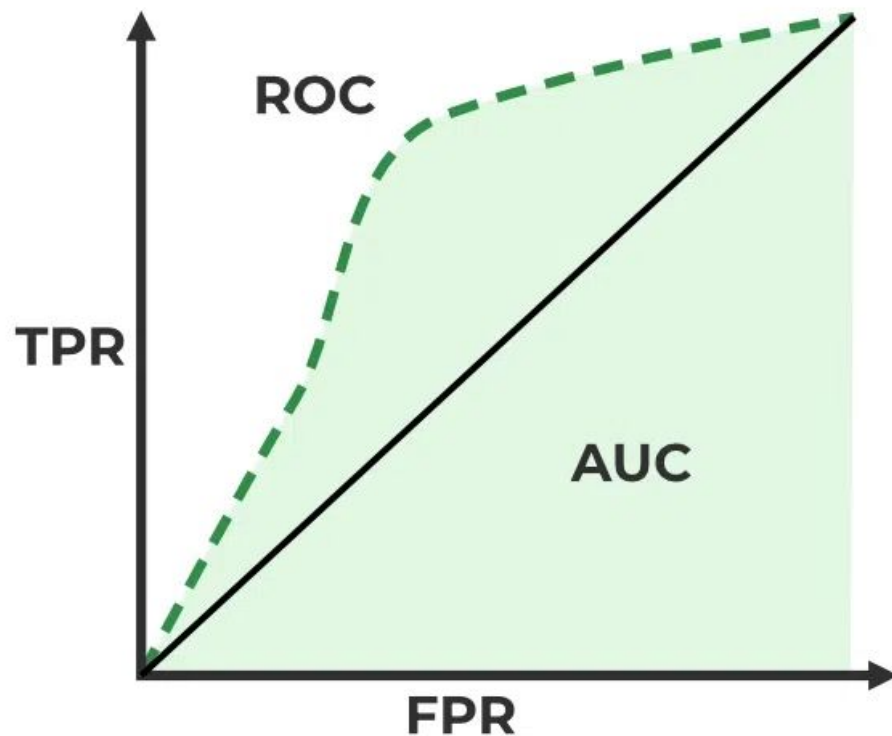
Recall:
ความน่าจะเป็นที่โมเดล
สามารถตรวจจับ
Positive จากจำนวน
ข้อมูล Positive
ทั้งหมด

Model Evaluation in Binary Classification

Predicted	Positive	Negative
	189	26
Actual		
Positive	51	234
Negative		

$$\begin{aligned}\text{F1-score} &= 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \\ &= 2 \times \frac{0.879 \times 0.788}{0.879 + 0.788} \\ &= 0.831\end{aligned}$$

Model Evaluation in Binary Classification

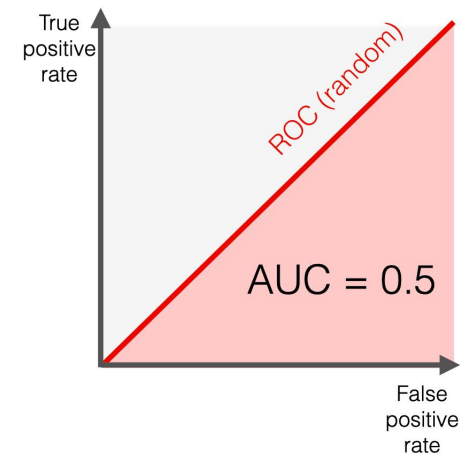
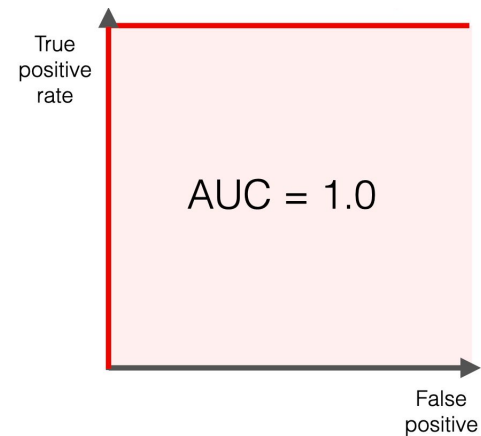
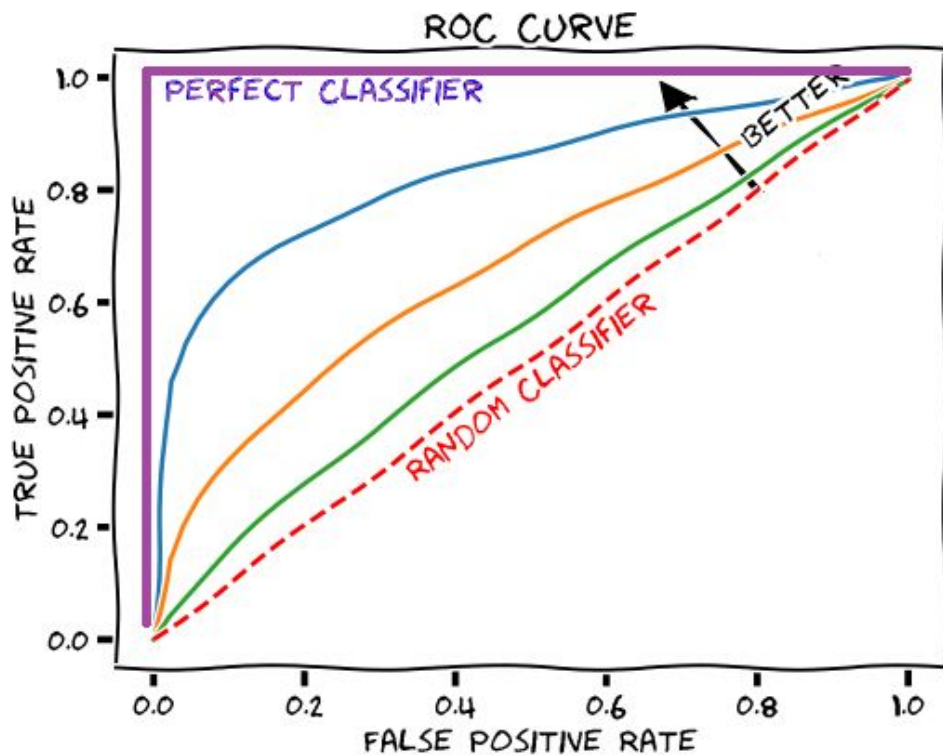


$$\text{AUC} = \int_0^1 \text{TPR}(t) d(\text{FPR}(t))$$

$$\text{TPR} = \frac{\text{TP}}{\text{TP} + \text{FN}}$$

$$\text{FPR} = \frac{\text{FP}}{\text{FP} + \text{TN}}$$

Model Evaluation in Binary Classification









Model Evaluation in Multiclass Classification

Predicted	TN	FN	TN
	FP	TP	FP
	TN	FN	TN
Actual			



Actual	TN	FP	TN
	FN	TP	FN
	TN	FP	TN
Predicted			

Model Evaluation in Multiclass Classification

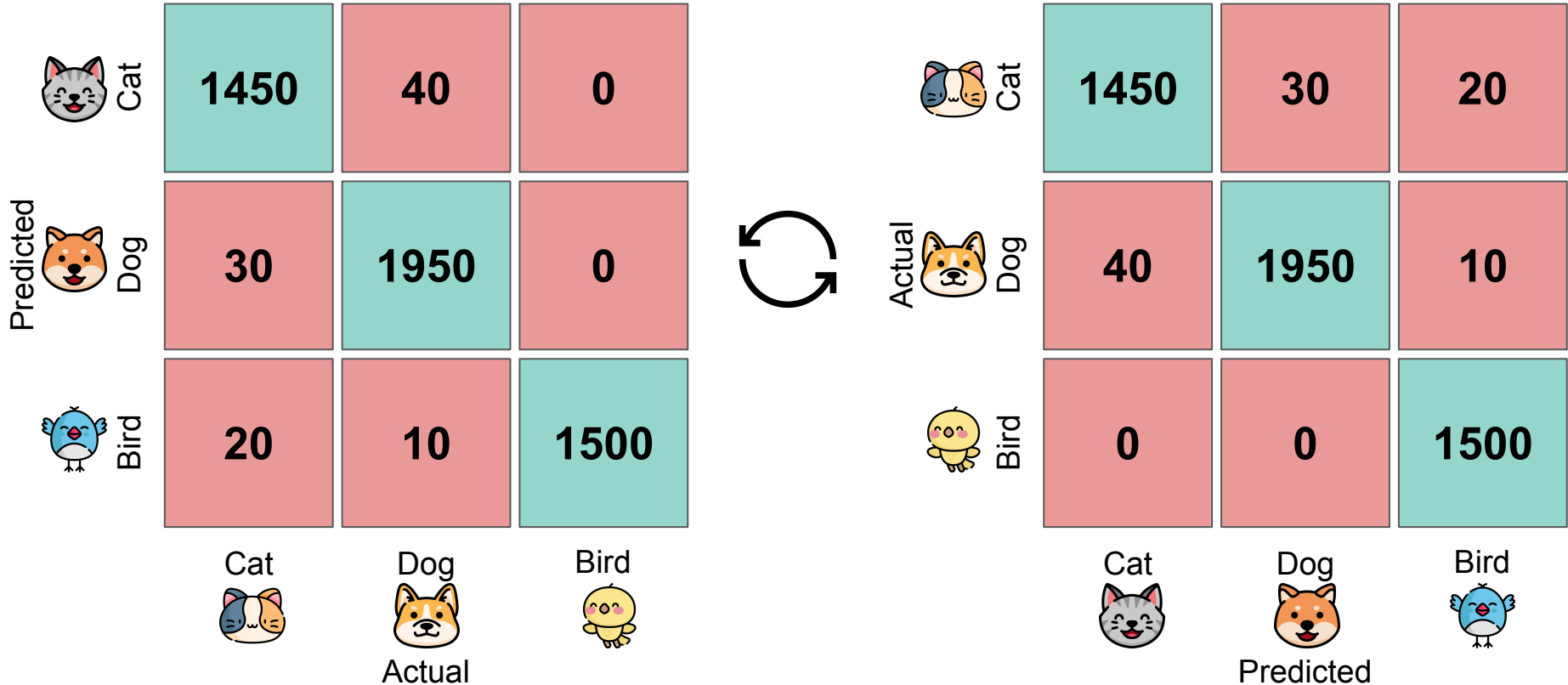
Predicted	 Cat	1450	40	0
	 Dog	30	1950	0
	 Bird	20	10	1500
		Cat 	Dog 	Bird 
		Actual		

ข้อมูลมีจำนวนทั้งหมด 5000 ภาพ และสามารถแบ่งออกเป็น 3 ประเภท ได้แก่ข้อมูล “Cat” มีจำนวน 1500 ภาพ, ข้อมูล “Cat” มีจำนวน 2000 ภาพ และข้อมูล “Bird” มีจำนวน 1500 ภาพ ผลลัพธ์ของตัวแบบที่เราสร้างขึ้นมีดังนี้







- จำแนกข้อมูล “Cat” ได้ถูกต้องเป็นจำนวน 1450 ภาพ
 - จำแนกข้อมูล “Cat” เป็น “Dog” จำนวน 30 ภาพ
 - จำแนกข้อมูล “Cat” เป็น “Bird” จำนวน 20 ภาพ
- จำแนกข้อมูล “Dog” ได้ถูกต้องเป็นจำนวน 1950 ภาพ
 - จำแนกข้อมูล “Dog” เป็น “Cat” จำนวน 40 ภาพ
 - จำแนกข้อมูล “Dog” เป็น “Bird” จำนวน 10 ภาพ
- จำแนกข้อมูล “Bird” ได้ถูกต้องทั้งหมด

จงทำการสร้าง Confusion Matrix สำหรับผลลัพธ์ดังกล่าว

Model Evaluation in Multiclass Classification



Model Evaluation in Multiclass Classification

Predicted	 Cat	1450	40	0
	 Dog	30	1950	0
	 Bird	20	10	1500
		Cat 	Dog 	Bird 
		Actual		

$$\begin{aligned}\text{Accuracy} &= \frac{TP_{\text{Cat}} + TP_{\text{Dog}} + TP_{\text{Bird}}}{N} \\ &= \frac{1450 + 1950 + 1500}{5000} \\ &= 0.98\end{aligned}$$

$$\begin{aligned}&= \frac{\sum_{i=1}^N \mathbb{I}(\hat{y}_i = y_i)}{N} \\ &= \frac{4900}{5000} = 0.98\end{aligned}$$







$$FP_{\text{Cat}} = 40$$

Predicted Cat but actual Dog or Bird

$$FN_{\text{Cat}} = 50$$

Actual Cat but predicted Dog or Bird







Model Evaluation in Multiclass Classification

Predicted	 Cat	1450	40	0
	 Dog	30	1950	0
	 Bird	20	10	1500
		Cat 	Dog 	Bird 
		Actual		

$$\begin{aligned} \text{Precision}_{\text{Cat}} &= \frac{TP_{\text{Cat}}}{TP_{\text{Cat}} + FP_{\text{Cat}}} \\ &= \frac{1450}{1450 + 40} \\ &= 0.97 \end{aligned}$$

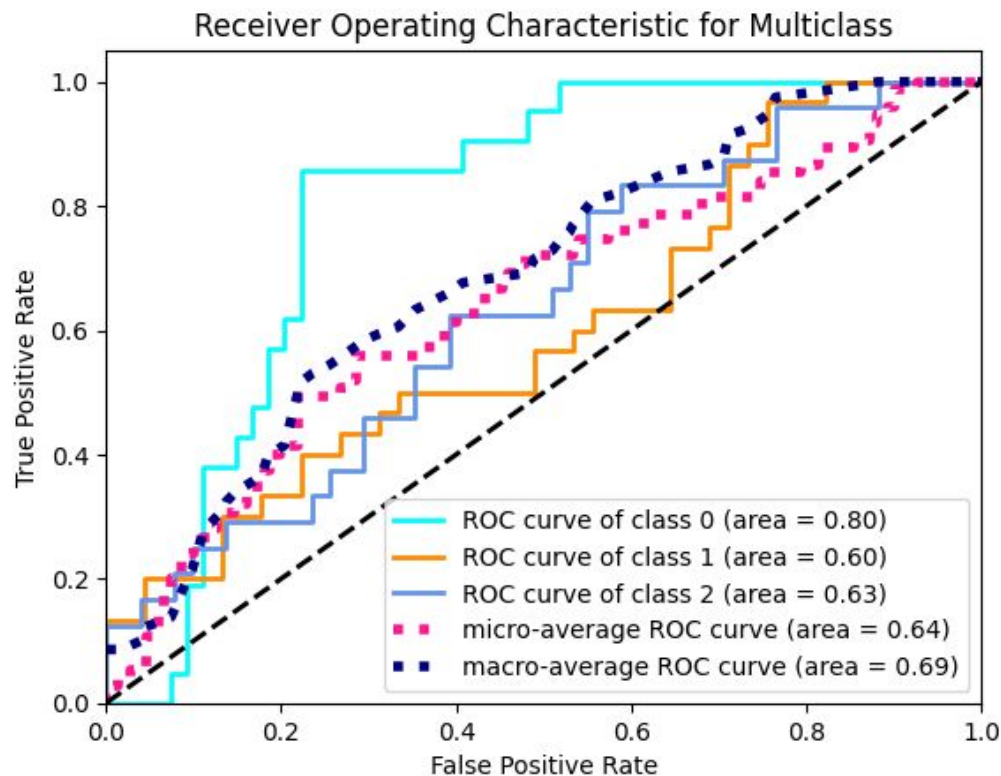
$$\begin{aligned} \text{Recall}_{\text{Cat}} &= \frac{TP_{\text{Cat}}}{TP_{\text{Cat}} + FN_{\text{Cat}}} \\ &= \frac{1450}{1450 + 30 + 20} \\ &= 0.96 \end{aligned}$$

Model Evaluation in Multiclass Classification

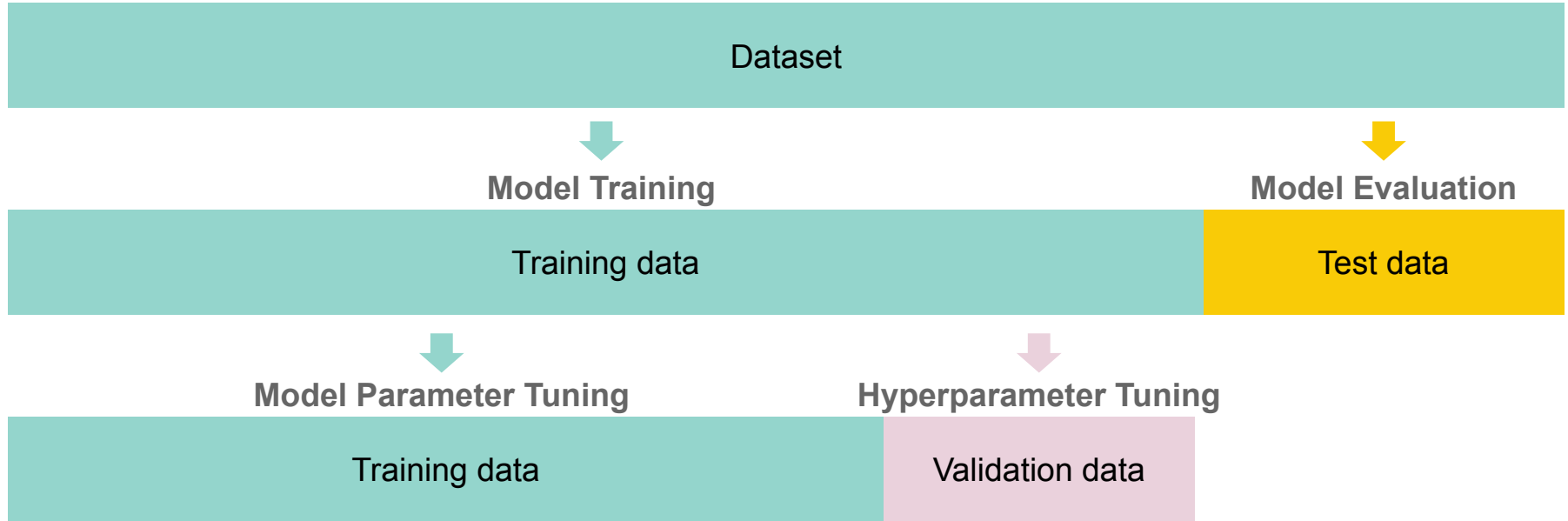
Predicted	 Cat	1450	40	0
	 Dog	30	1950	0
	 Bird	20	10	1500
		Cat 	Dog 	Bird 
		Actual		

$$\begin{aligned}\text{F1-score}_{\text{Cat}} &= 2 \times \frac{\text{Precision}_{\text{Cat}} \times \text{Recall}_{\text{Cat}}}{\text{Precision}_{\text{Cat}} + \text{Recall}_{\text{Cat}}} \\ &= 2 \times \frac{0.97 \times 0.96}{0.97 + 0.96} \\ &= 0.964\end{aligned}$$

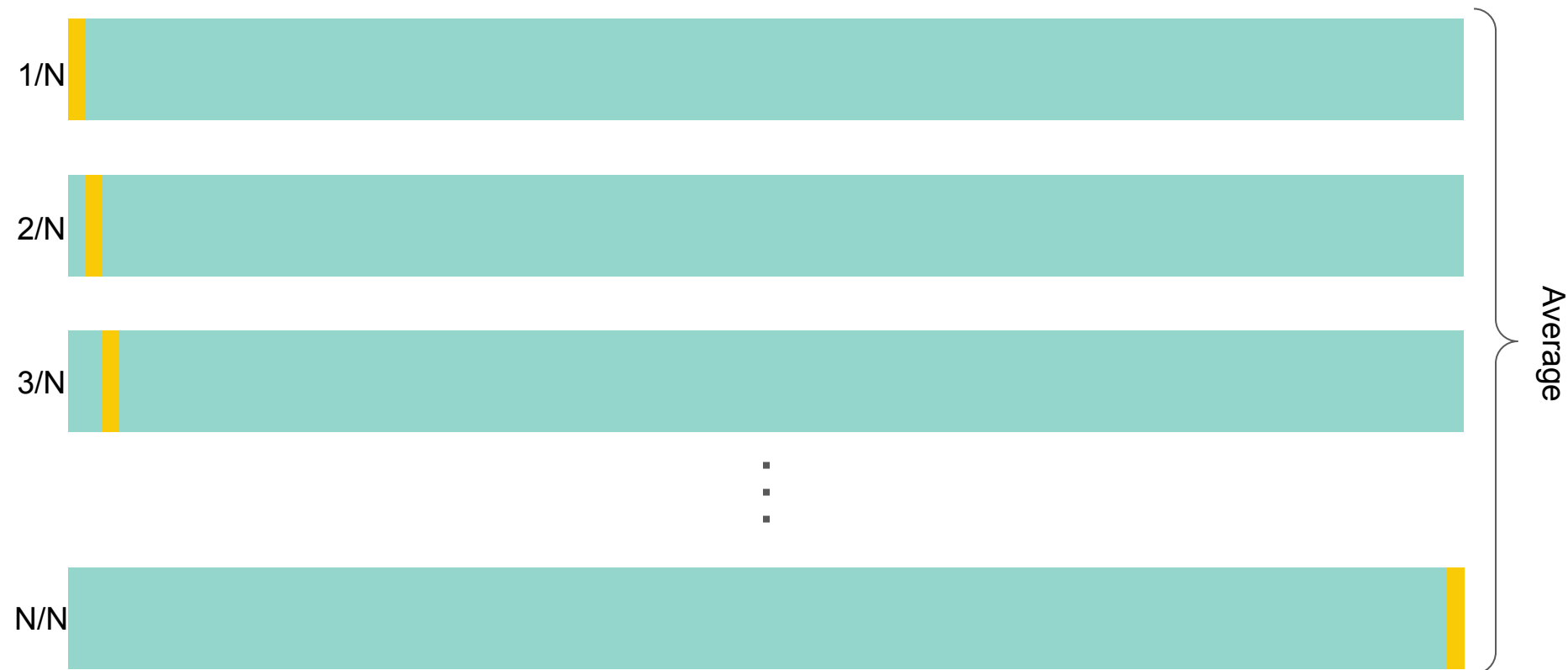
Model Evaluation in Multiclass Classification



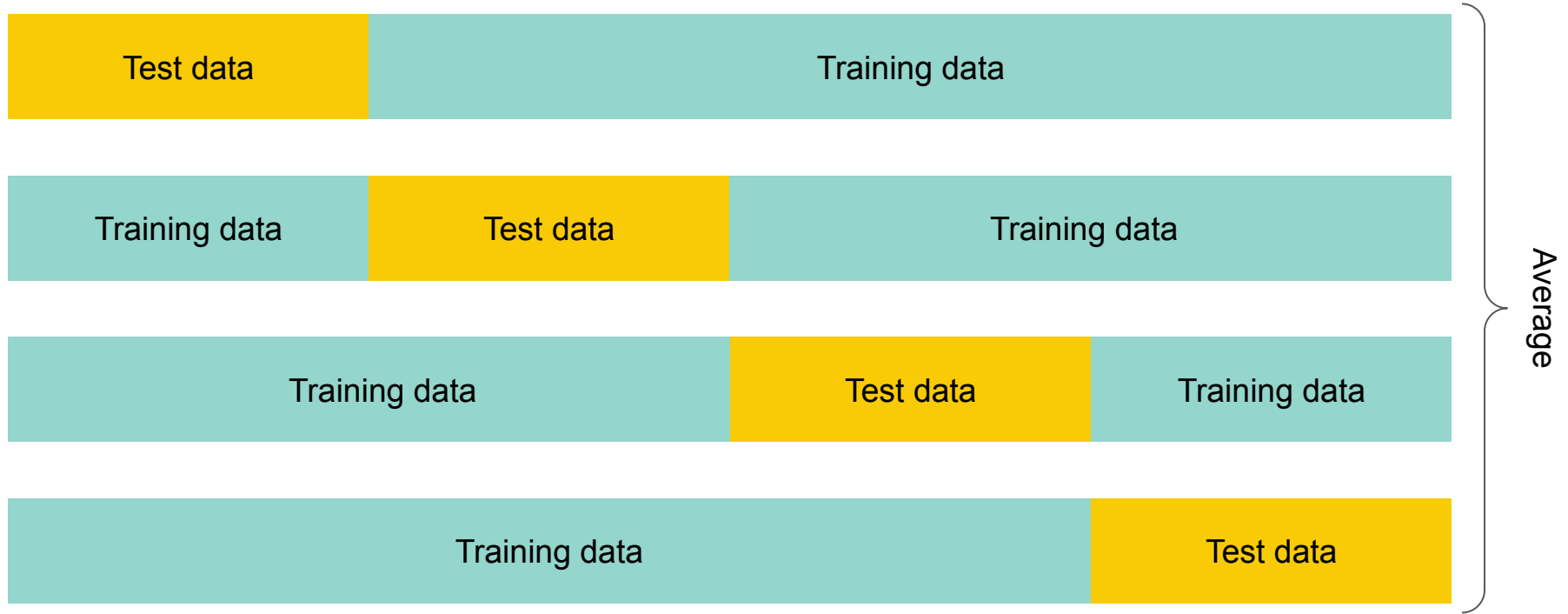
Cross-Validation Techniques: Holdout



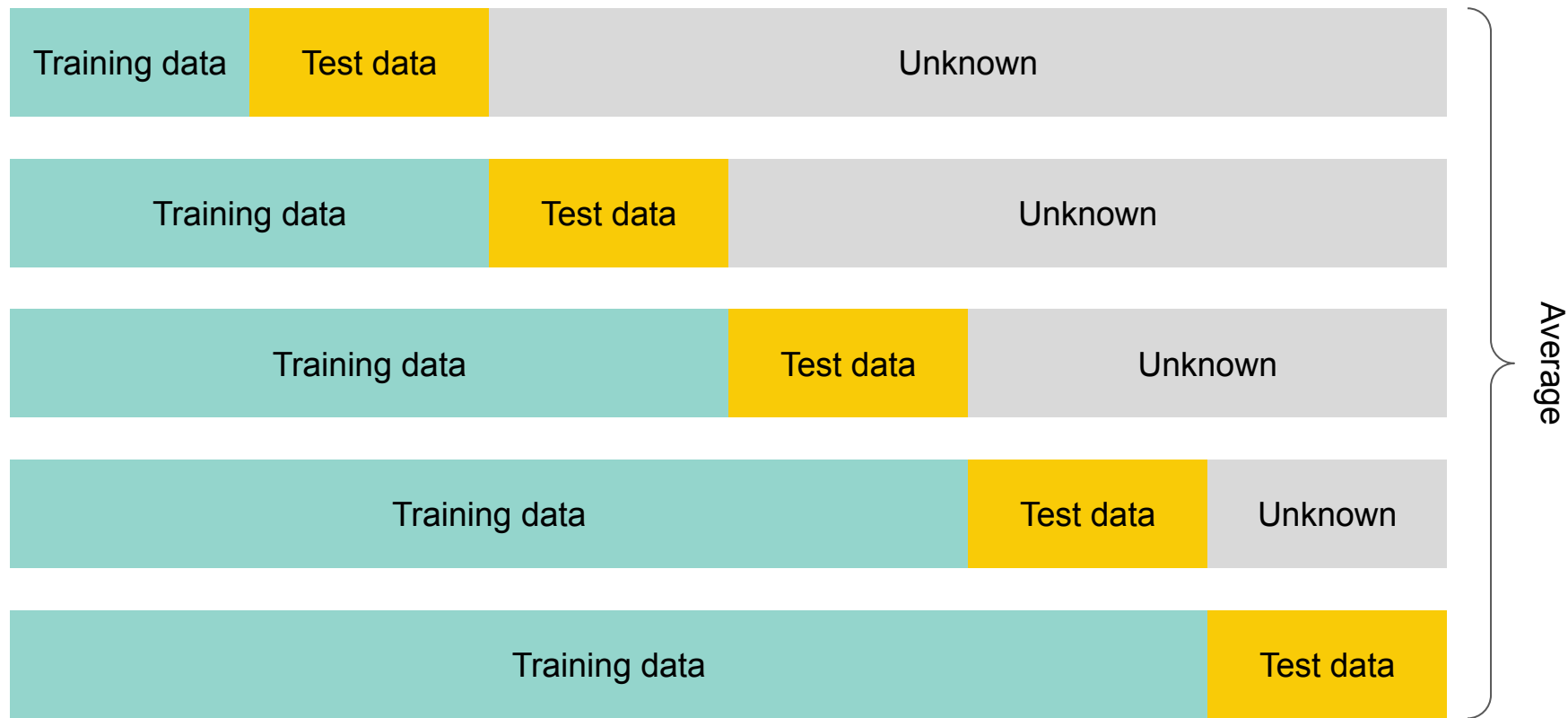
Cross-Validation Techniques: Leave One Out



Cross-Validation Techniques: K-Fold



Cross-Validation Techniques: K-Fold (Time-Series)



Workshop

หากเราทำการสร้างตัวแบบสำหรับการจำแนกข้อมูลของการตัดสินใจซื้อรถของลูกค้า โดยข้อมูลฝึกดังกล่าวมีจำนวนทั้งหมด 1000 ข้อมูล และสามารถแบ่งออกเป็น 2 ประเภท ได้แก่ ข้อมูล “ซื้อ” มีจำนวน 750 ข้อมูล และข้อมูล “ไม่ซื้อ” ผลปรากฏว่าตัวแบบที่เราสร้างขึ้นสามารถทำการจำแนกข้อมูล “ซื้อ” ได้ถูกต้องเป็นจำนวน 680 ข้อมูล แต่จำแนกข้อมูล “ไม่ซื้อ” ผิดเป็นจำนวน 34 ข้อมูล

จงทำการสร้าง Confusion Matrix สำหรับผลลัพธ์ดังกล่าว พร้อมทั้งทำการคำนวณหาค่า Accuracy, Precision, Recall และ F1-score

Workshop

จงทำการคำนวณหาค่า Precision, Recall และ F1-score สำหรับคลาส “Dog” และ “Bird”

Predicted	Cat	1450	40	0
	Dog	30	1950	0
	Bird	20	10	1500
		Cat	Dog	Bird
		Actual		

Workshop

จงคำนวณหาค่า MSE, MAE, MAPE และ R^2

$$\hat{\mathbf{y}} = [0.0, 0.2, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.1]$$

$$\mathbf{y} = [0.1, 0.3, 0.5, 0.6, 0.5, 0.55, 0.75, 0.89, 0.99]$$