

Text Classification Using Machine Learning and Deep Learning

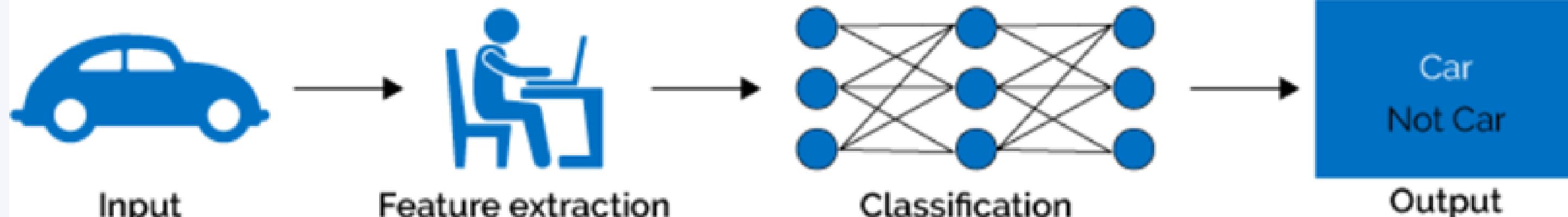
Topic 6

Dwi Intan Af'ida, S.T., M.Kom

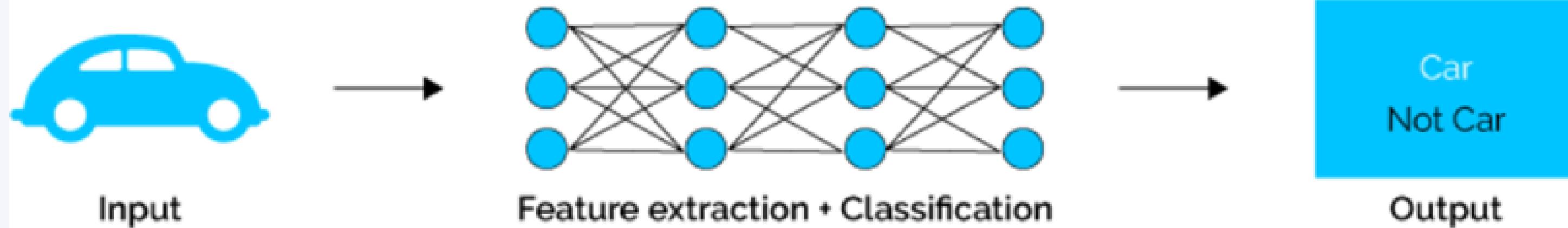


ML vs DL

Machine Learning



Deep Learning



ML vs DL

- 
- Machine Learning merupakan cabang ilmu komputer dengan fokus pada pengembangan sebuah sistem yang mampu belajar sendiri tanpa harus berulang kali di program oleh manusia.
 - Peran awal data sangat penting sebagai langkah awal pada Machine Learning untuk menghasilkan output.
 - Hal ini digunakan sebagai latihan atau uji coba awal dari Machine Learning.
 - Setelah melewati uji coba awal, Machine Learning akan dapat menyelesaikan masalah tanpa diprogram secara eksplisit.
 - Deep Learning, di lain sisi, merupakan salah satu metode implementasi dari Machine Learning yang bertujuan untuk meniru cara kerja otak manusia menggunakan Artificial Neural Network atau jaringan nalar buatan.
 - Deep Learning dengan sejumlah algoritmanya sebagai "neuron" akan bekerja sama dalam menentukan dan mencerna karakteristik-karakteristik tertentu pada suatu rangkaian data.
 - Program dalam Deep Learning biasanya menggunakan kapabilitas yang lebih kompleks dalam mempelajari, mencerna, dan juga mengklasifikasikan data

ML vs DL

Salah satu perbedaan utama antara Machine Learning dan Deep Learning adalah performanya ketika jumlah data terus meningkat dan bagaimana menyelesaikan suatu masalah.

- algoritma Machine Learning mampu mengolah data dalam jumlah yang lebih kecil.
- Dan untuk menyelesaikan masalah disarankan memecahkannya menjadi beberapa bagian agar dapat diselesaikan secara terpisah, dan penyelesaiannya digabungkan guna mendapatkan hasil yang utuh.
- Algoritma Deep Learning digunakan untuk membuat jaringan syaraf buatan yang tidak mampu mengolah data dalam jumlah kecil secara maksimal.
- Hal ini karena algoritma Deep Learning membutuhkan data dalam jumlah banyak dan mampu menyelesaikan masalah secara keseluruhan dari awal hingga akhir tanpa perlu memisahkannya menjadi beberapa bagian.

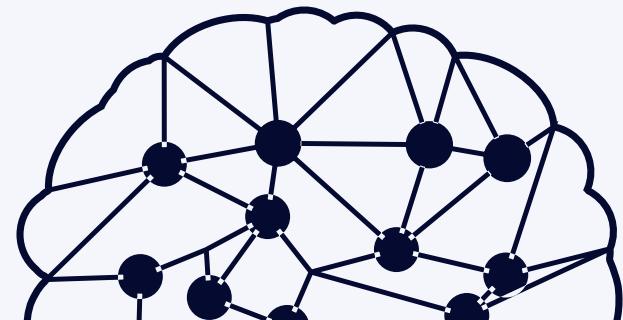
Building Model of ML-DL

01 Split Dataset

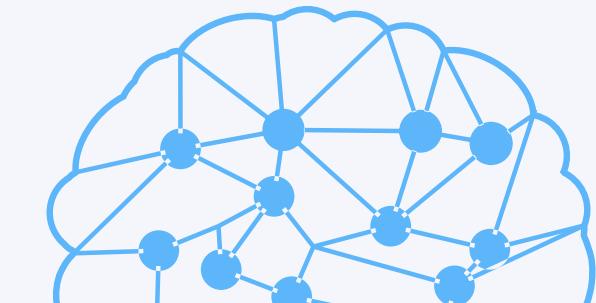
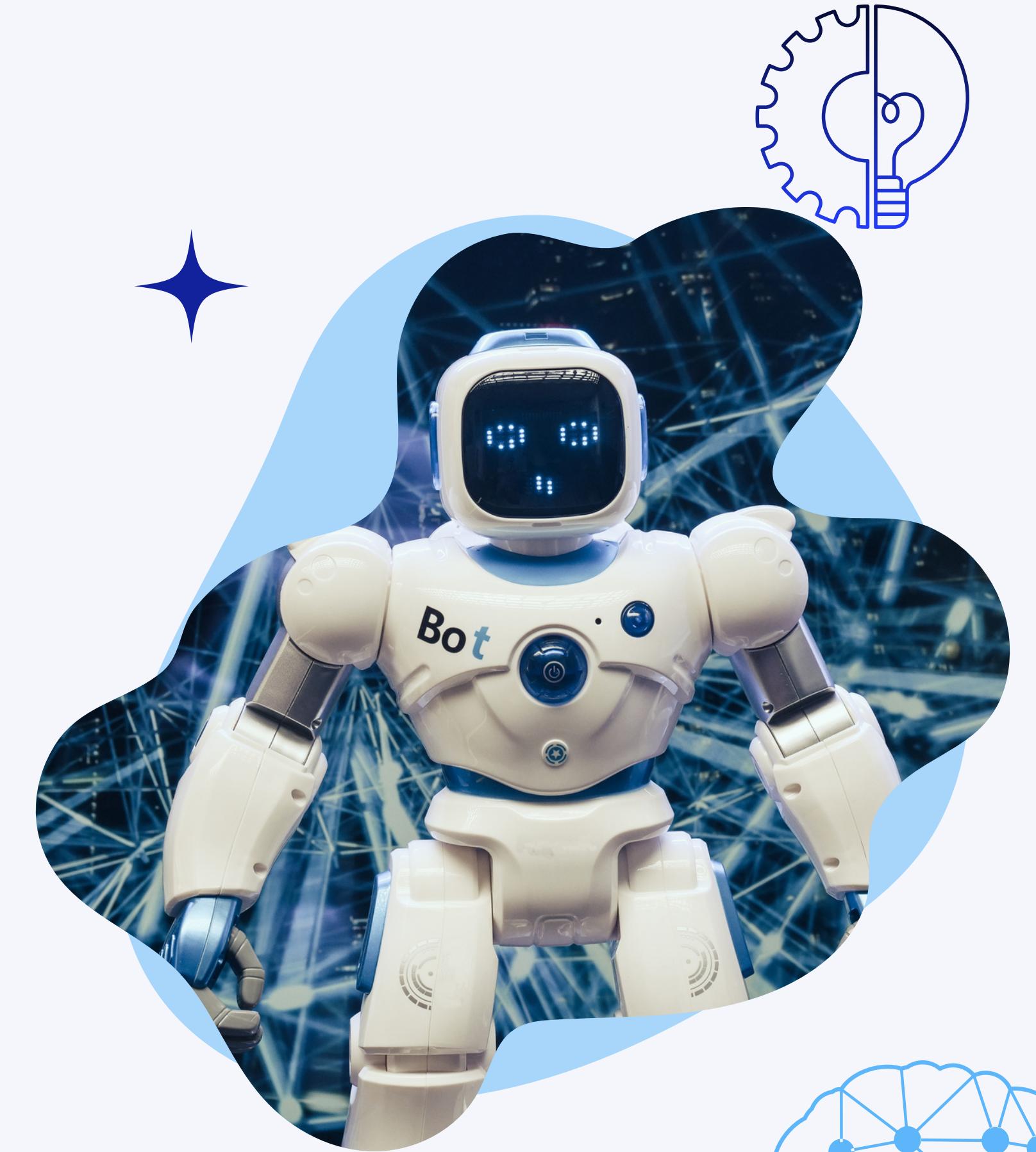
02 Traning

03 Testing

03 Evaluation Method



Split Data



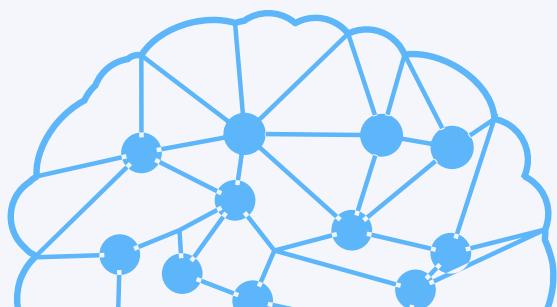
Split Data

Data Training & Data Testing

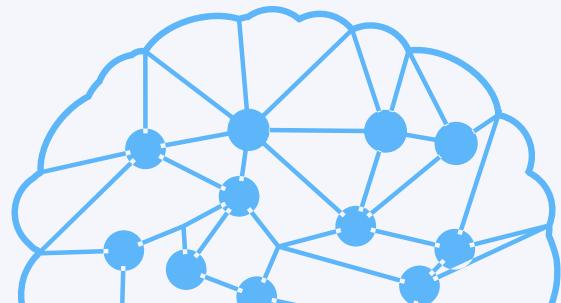
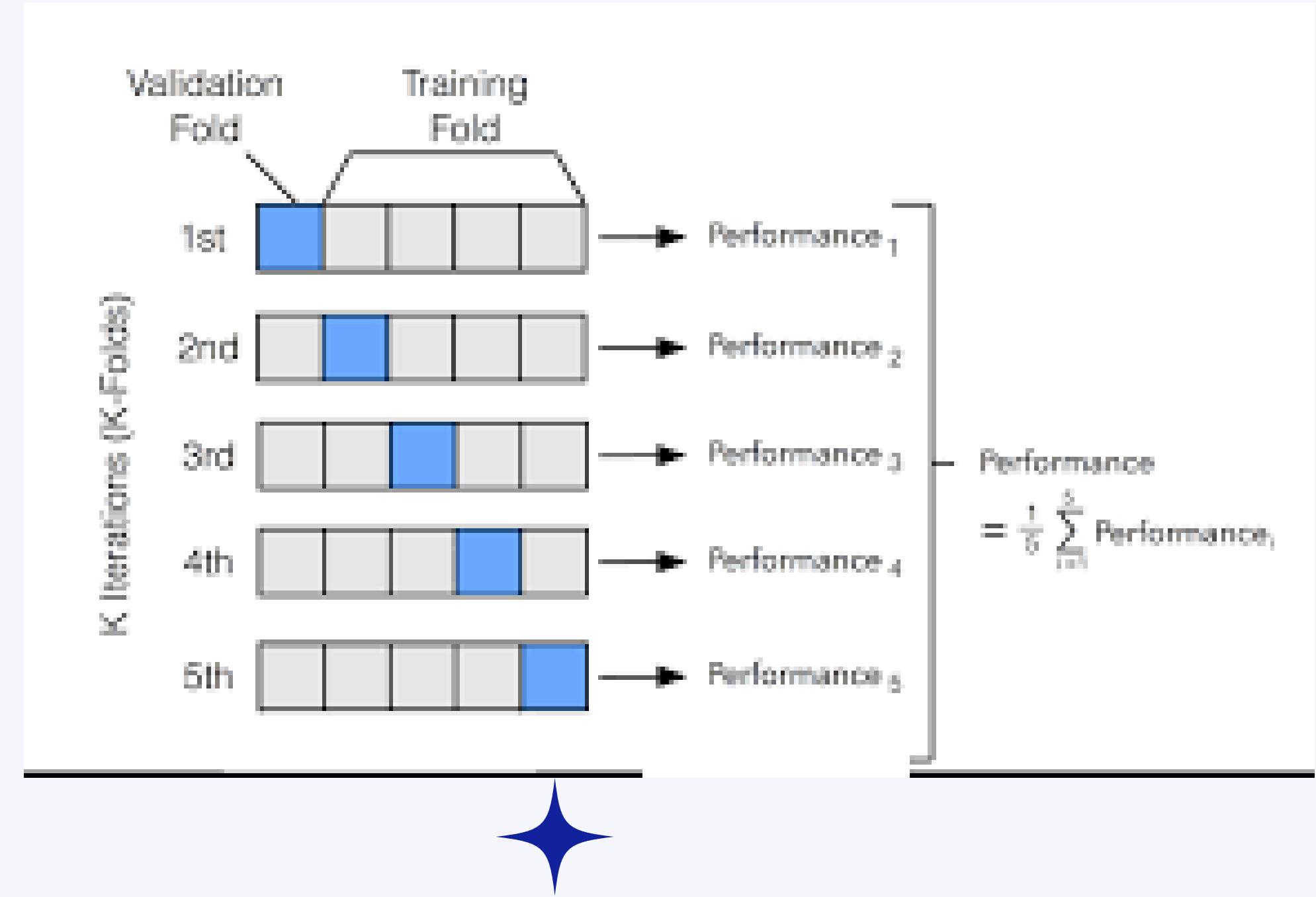
01 70 : 30

02 80 : 20

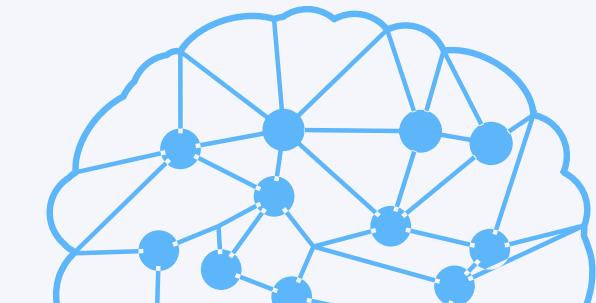
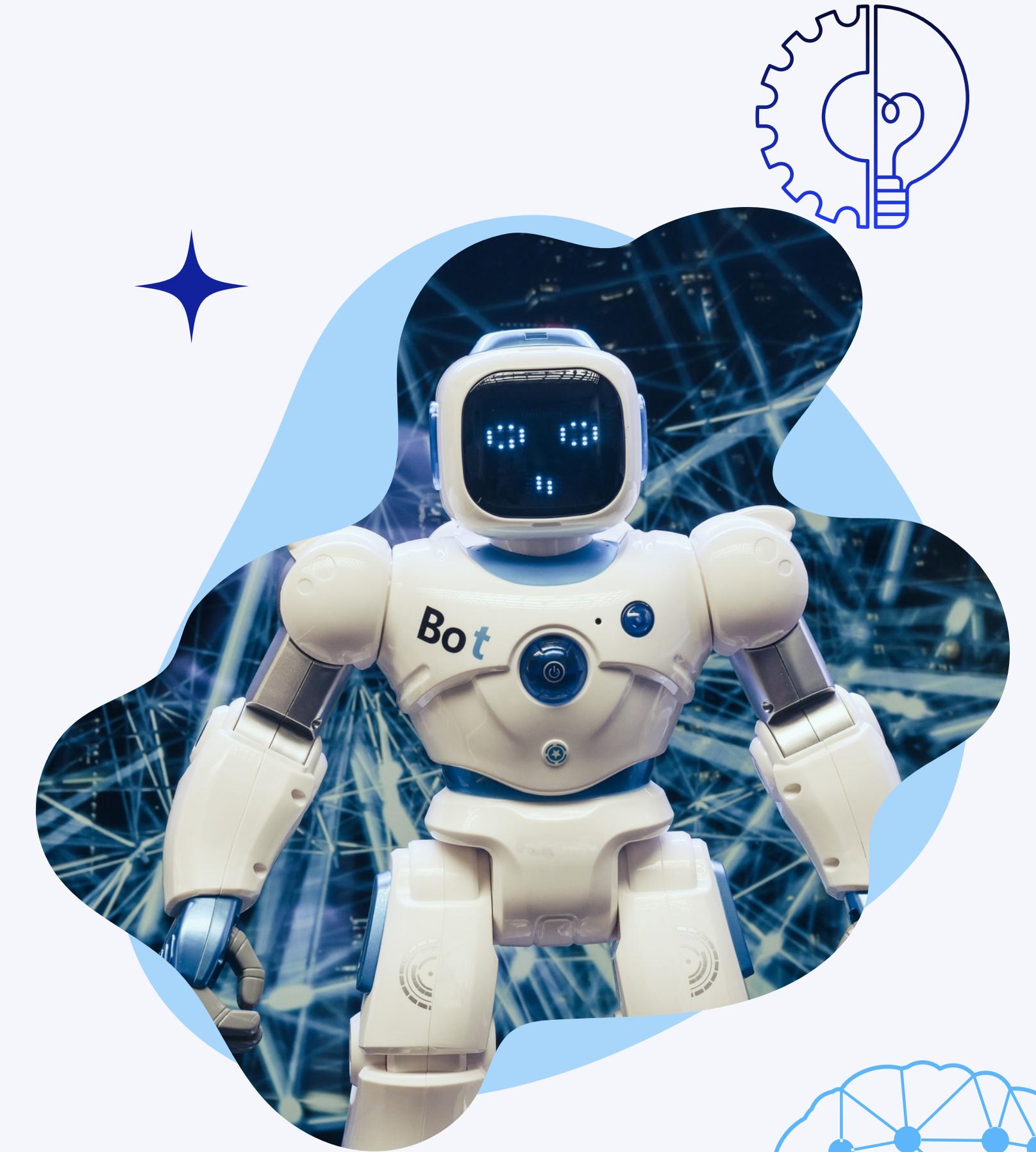
03 K-Fold Cross Validation



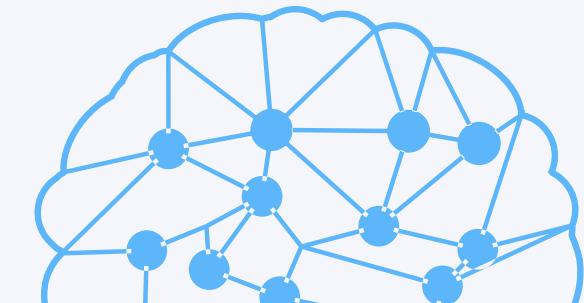
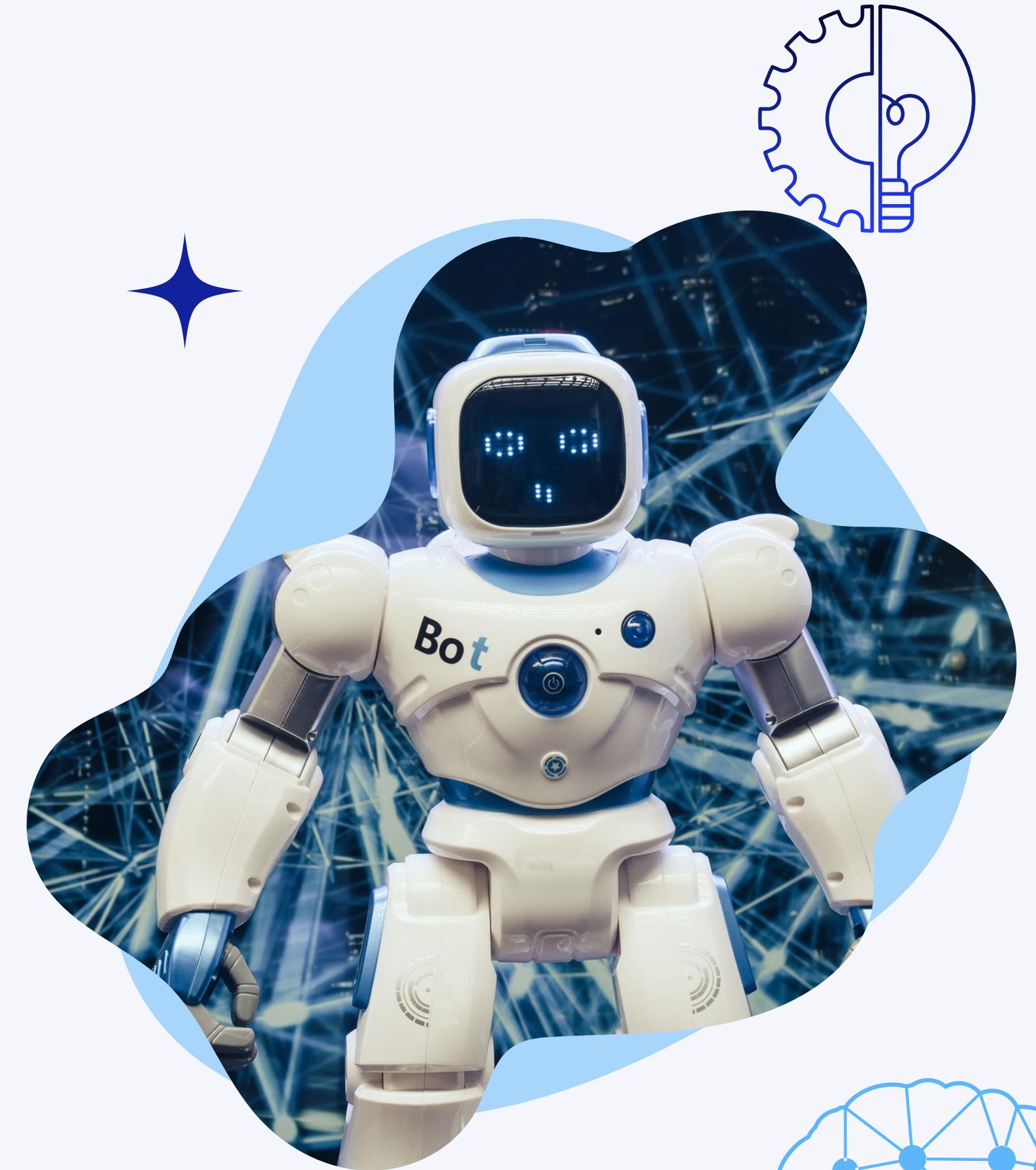
K-Fold



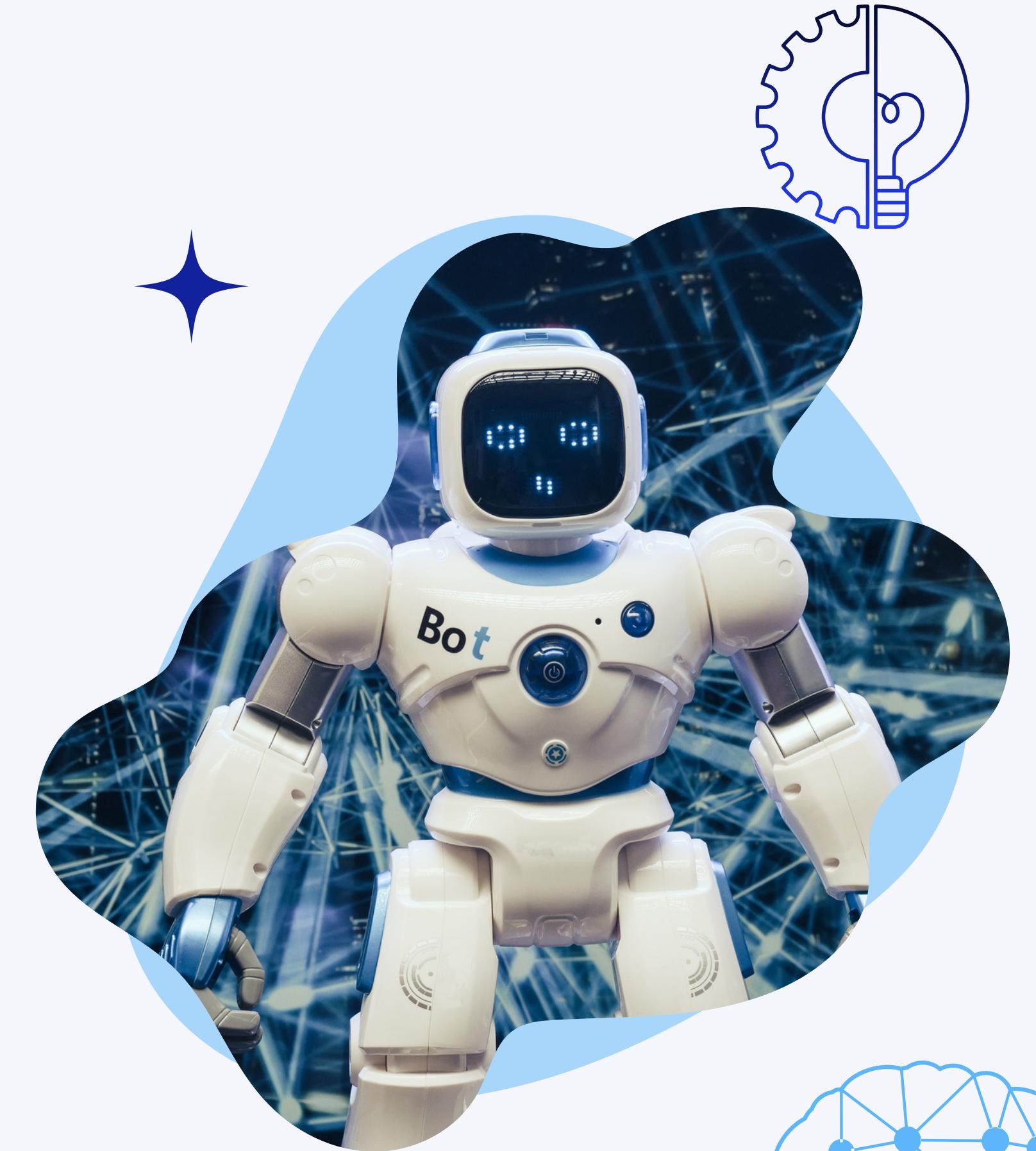
Training



Testing



Evaluation



Confussion Matrix

		CONDITION determined by "Gold Standard"			
		CONDITION POS	CONDITION NEG	PREVALENCE $\frac{\text{CONDITION POS}}{\text{TOTAL POPULATION}}$	
TEST OUT- COME	TEST POS	True Pos TP	Type I Error False Pos FP	Precision Pos Predictive Value $\text{PPV} = \frac{\text{TP}}{\text{TEST P}}$	False Discovery Rate FDR = $\frac{\text{FP}}{\text{TEST P}}$
	TEST NEG	Type II Error False Neg FN	True Neg TN	False Omission Rate FOR = $\frac{\text{FN}}{\text{TEST N}}$	Neg Predictive Value NPV = $\frac{\text{TN}}{\text{TEST N}}$
ACCURACY ACC $\text{ACC} = \frac{\text{TP} + \text{TN}}{\text{TOT POP}}$	Sensitivity (SN), Recall Total Pos Rate TPR $\text{TPR} = \frac{\text{TP}}{\text{CONDITION POS}}$	Fall-Out False Pos Rate FPR $\text{FPR} = \frac{\text{FP}}{\text{CONDITION NEG}}$	Pos Likelihood Ratio LR^+ $\text{LR}^+ = \frac{\text{TPR}}{\text{FPR}}$	Diagnostic Odds Ratio DOR $\text{DOR} = \frac{\text{LR}^+}{\text{LR}^-}$	
	Miss Rate False Neg Rate FNR $\text{FNR} = \frac{\text{FN}}{\text{CONDITION POS}}$	Specificity (SPC) True Neg Rate TNR $\text{TNR} = \frac{\text{TN}}{\text{CONDITION NEG}}$	Neg Likelihood Ratio LR^- $\text{LR}^- = \frac{\text{TNR}}{\text{FNR}}$		



Thank You

credit: canva.com

