



**Google Developer Student Club**  
Indian Institute Of Information Technology Vadodara

# Classification Algorithms

**Aman Gangwar**  
**Data Science Lead**  
**GDSC IIITV**

# Classification Algorithms

Supervised

Unsupervised



# Supervised Learning

# Supervision?

# Logistic Regression

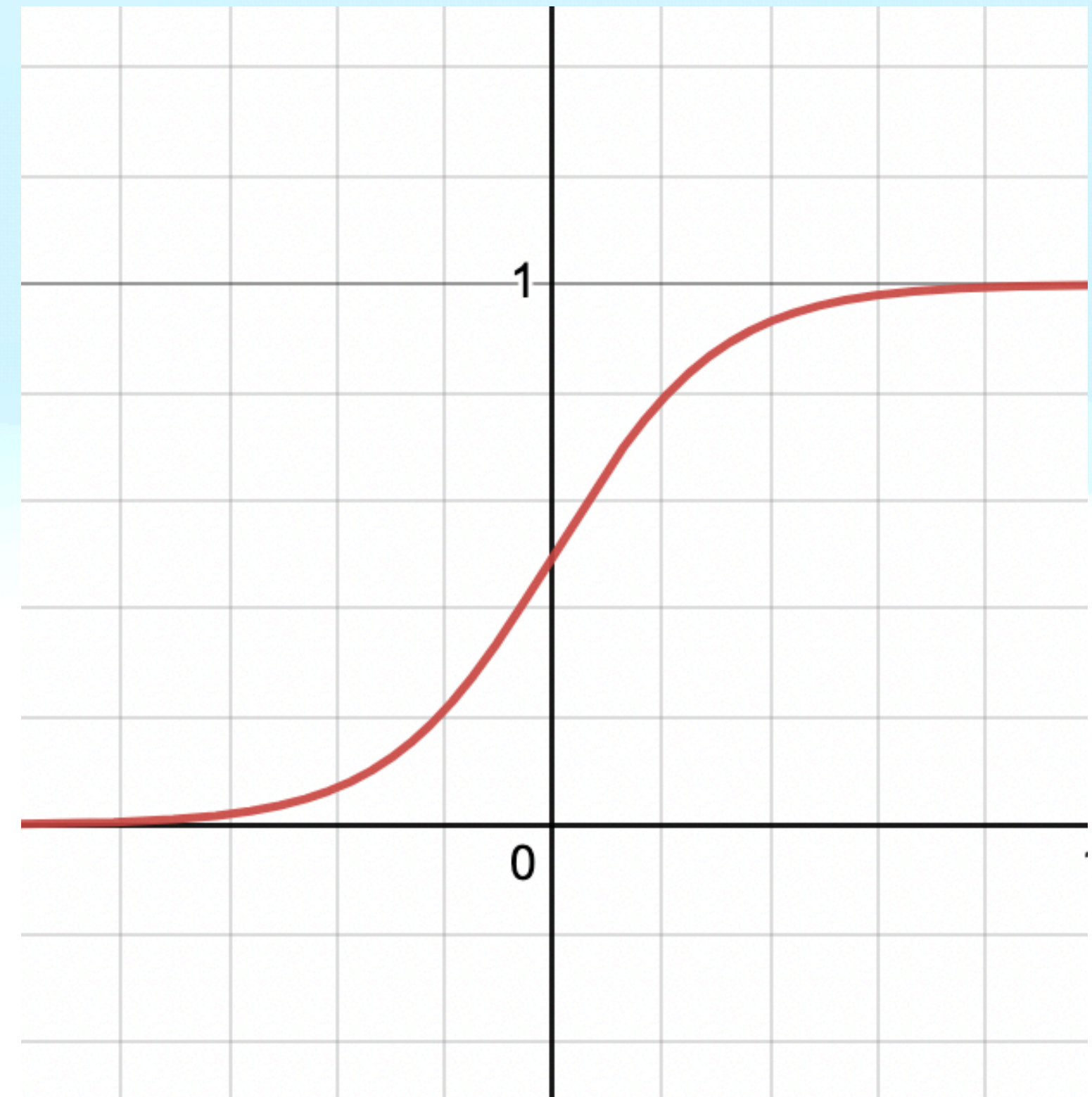


# Logistic Regression?



# Sigmoid Function

$$f(x) = \frac{1}{1 + e^{-x}}$$





# Multiple Variables

$$P(y = 1 \mid x_1, \dots, x_n) = \frac{1}{1 + e^{-(b_1x_1 + b_2x_2 + \dots + b_kx_k + a)}}$$



# Loss Function

$$loss = -y\log(y_p) - (1 - y)\log(1 - y_p)$$

# Decision Trees



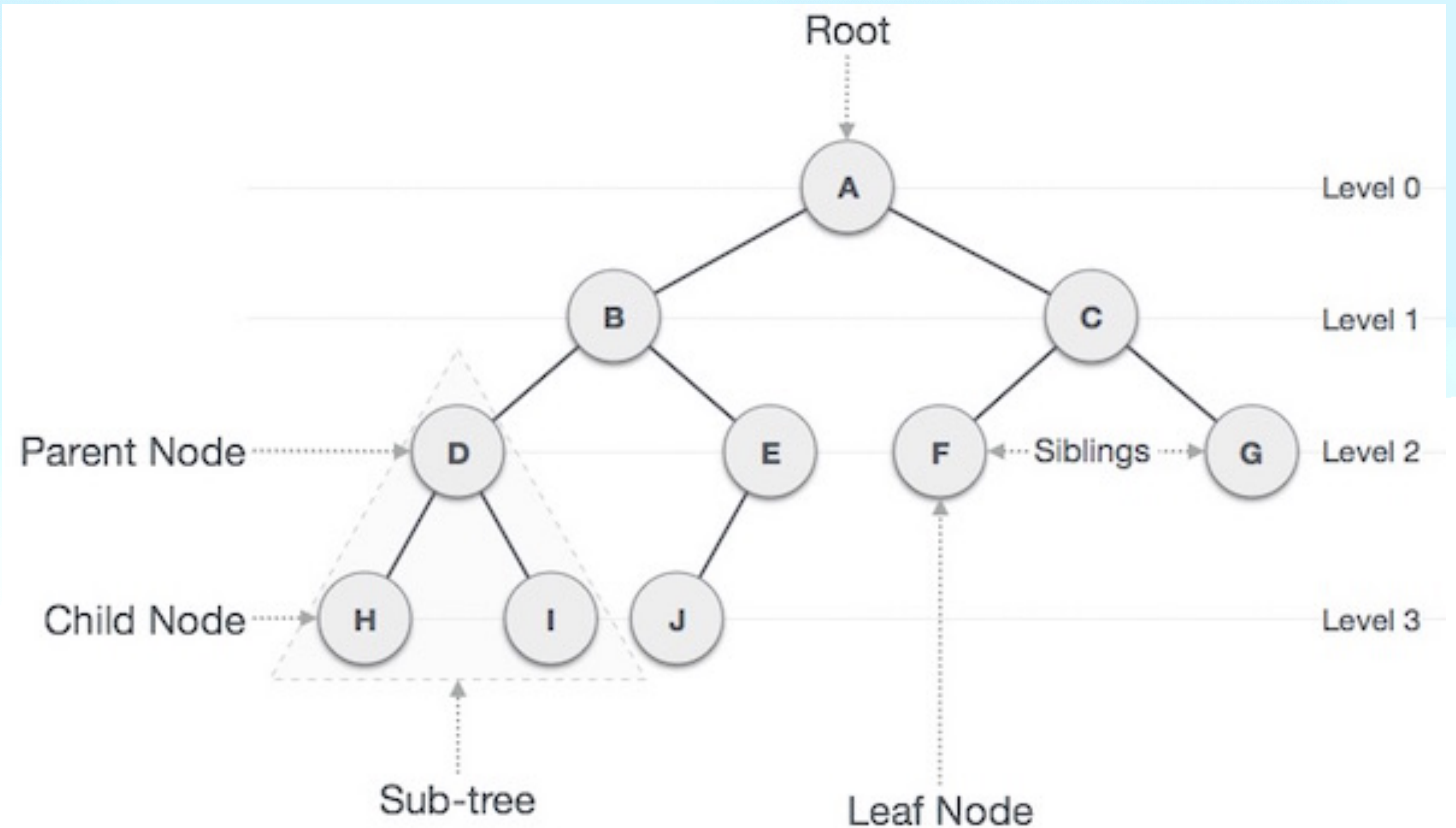
# Decision Trees



# Tree

## Data Structure

- Root
- Node
- Child





**data is classified in the form of  
trees.**

**want more?**



# Naïve Bayes

# Support Vector Machines



# Metrics

# Confusion Matrix

Predicted Values

		Actual Values	
		Positive (1)	Negative (0)
Predicted Values	Positive (1)	TP	FP
	Negative (0)	FN	TN



# Precision

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

Predicted Values

Actual Values	
	Positive (1)    Negative (0)
Positive (1)	TP                  FP
Negative (0)	FN                  TN

# Recall

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

Predicted Values

		Actual Values	
		Positive (1)	Negative (0)
Predicted Values	Positive (1)	TP	FP
	Negative (0)	FN	TN

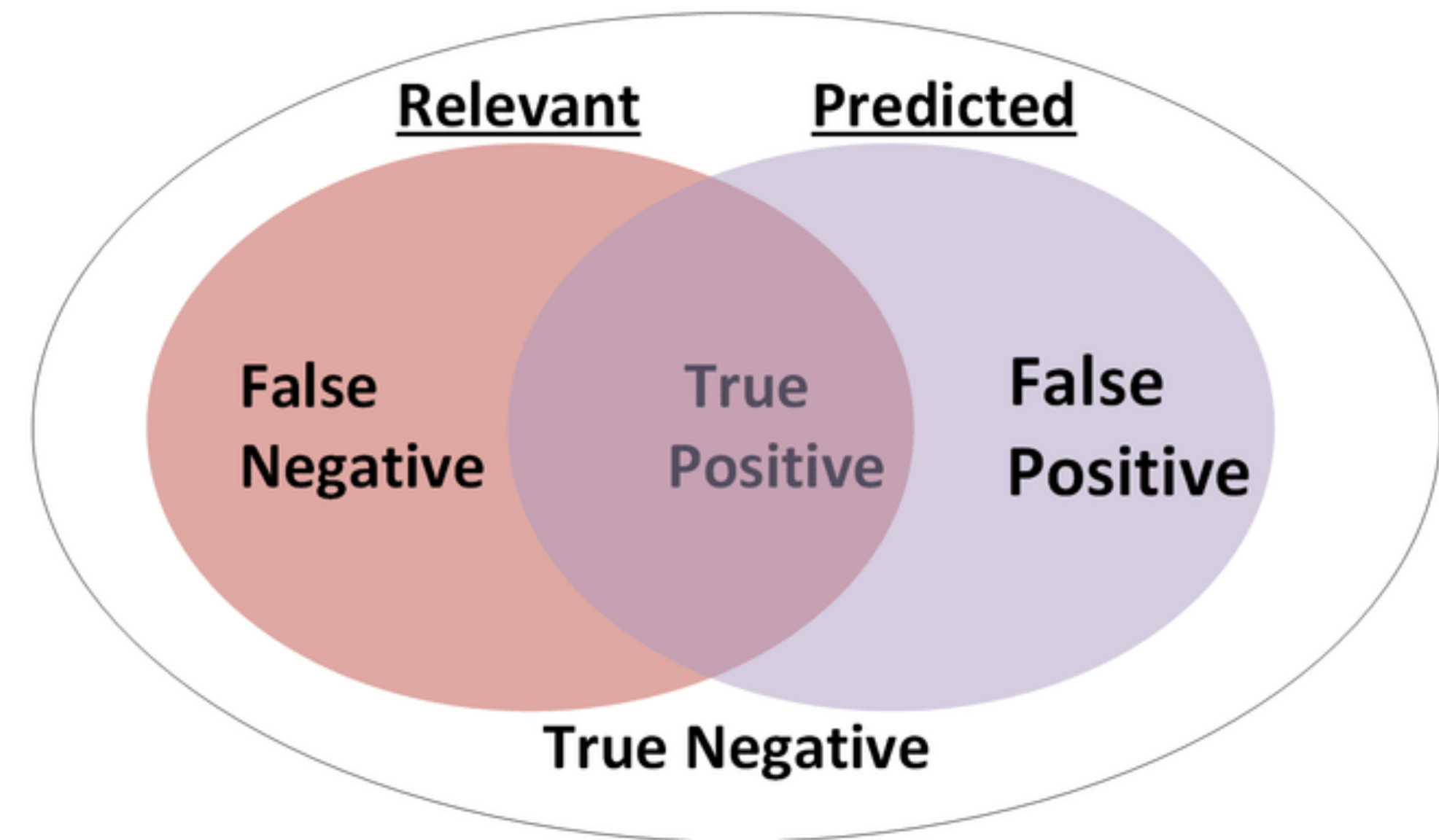


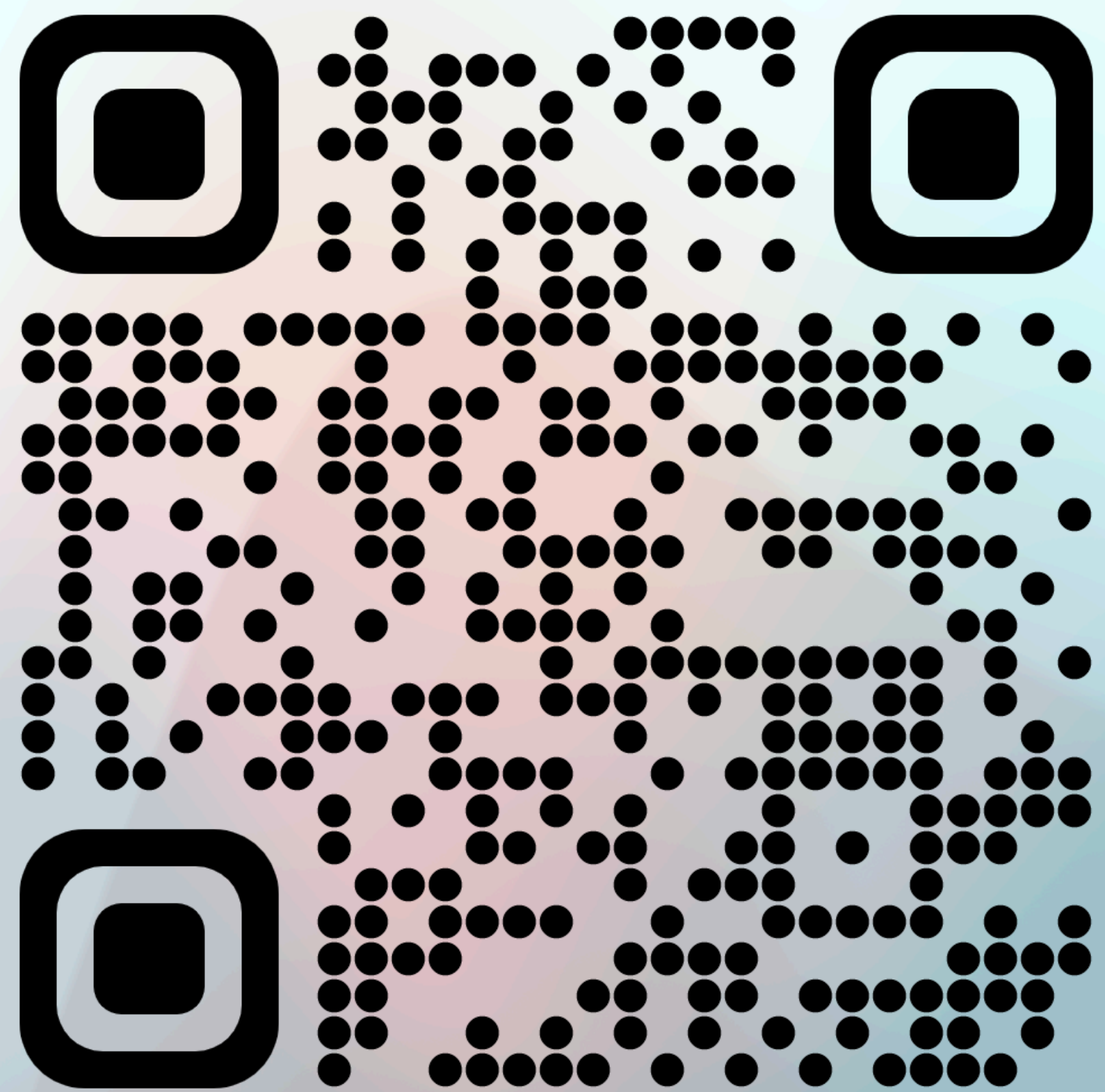
# Precision & Recall

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

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

	Relevant	Not relevant
Predicted	True Positive (TP)	False Positive (FP)
Not Predicted	False Negative (FN)	True Negative (TN)







# Unsupervised Learning

# K-means Clustering



धन्यवाद

Thank You

ありがとう