

# HỌC MÁY ỨNG DỤNG

(Applied Machine Learning)

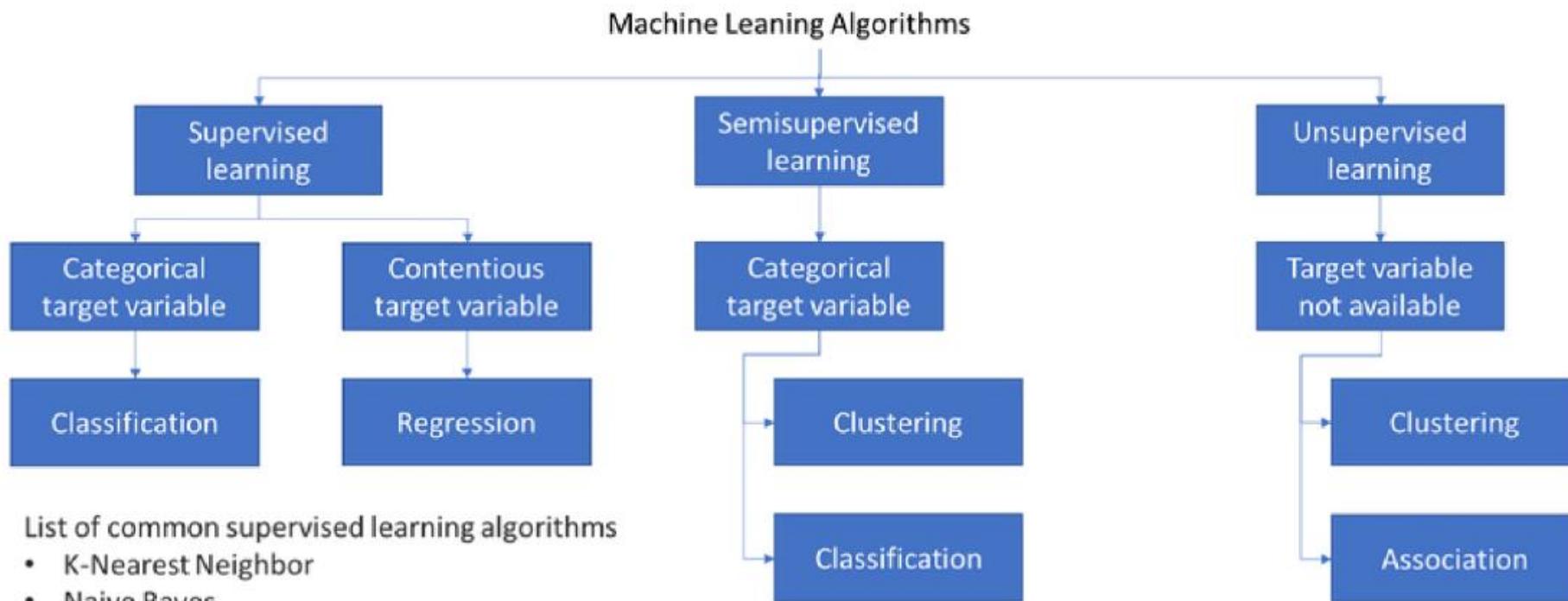
## CHƯƠNG 2: HỌC TƯƠNG ĐỒNG (SIMILAR BASED LEARNING)



GVGD: PHAN HỒ VIẾT TRƯỜNG



# GIỚI THIỆU HỌC MÁY



List of common supervised learning algorithms

- K-Nearest Neighbor
- Naive Bayes
- Decision Trees
- Linear Regression
- Support Vector Machines (SVM)
- Neural Networks
- Classification and Regression Trees
- Gradient Boosted Regression Tree
- Perceptron Back-Propagation
- Random Forest

List of common semi supervised learning algorithms:

- Linear Regression
- Logistic Regression

List of common unsupervised learning algorithms:

- k-means clustering and classification
- Association Rules



# GIỚI THIỆU HỌC TƯƠNG ĐỒNG

- Học tương đồng là kỹ thuật học giám sát.
- Dự báo nhãn (label) của một trường hợp test (instance) bằng đo lường tương đồng giữa trường hợp test với các trường hợp khác trong tập huấn luyện.
- Cải tiến hiệu suất việc phân lớp
- Ứng dụng trong xử lý ảnh, phân lớp văn bản, nhận dạng mẫu ...



# GIỚI THIỆU HỌC TƯƠNG ĐỒNG

- Học tương đồng không xây dựng mô hình tổng quát.
- Lưu trữ tập huấn luyện và dùng khi cần phân lớp một trường hợp chưa biết.
- **Ưu điểm:**
  - Hữu ích khi tập huấn luyện không có sẵn, dùng cho việc tăng cường dữ liệu.
- **Nhược điểm:**
  - Cần bộ nhớ lớn để lưu trữ



# GIỚI THIỆU HỌC TƯƠNG ĐỒNG

- Học tương đồng không xây dựng mô hình tổng quát → còn gọi lazy learning
- Lưu trữ tập huấn luyện và dùng khi cần phân lớp một trường hợp chưa biết.
- Ưu điểm:
  - Hữu ích khi tập huấn luyện không có sẵn, dùng cho việc tăng cường dữ liệu.
- Nhược điểm:
  - Cần bộ nhớ lớn để lưu trữ



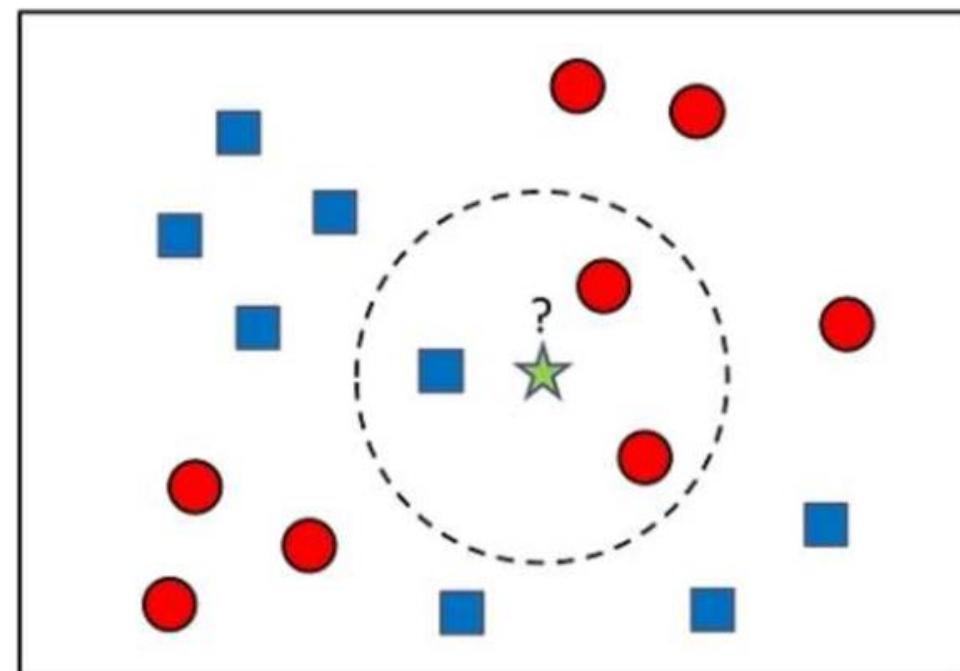
# HỌC TƯƠNG ĐỒNG <> HỌC MÔ HÌNH

Học tương đồng	Học mô hình
Lazy learner	Eager learner
Xử lý tập huấn luyện khi dự báo 1 trường hợp mới (bước Test)	Xử lý tập huấn luyện trong quá trình học (bước Train)
Không xây dựng model từ tập huấn luyện	Xây dựng model từ tập huấn luyện
Dự báo lớp trực tiếp từ tập train	Dự báo lớp từ mô hình đã xây
Tốc độ dự báo chậm	Tốc độ dự báo nhanh



# GIẢI THUẬT K-NEAREST NEIGHBOR (K-NN)

- Là phương pháp phi tham số dùng cho bài toán phân lớp và hồi quy.
- Dự báo lớp cho 1 trường hợp dựa vào k mẫu trong tập train gần với nó.
- Xác suất cao nhất các nhãn trong tập k mẫu → là nhãn của trường hợp





# GIẢI THUẬT K-NEAREST NEIGHBOR (K-NN)

---

**Algorithm** The  $k$ -nearest neighbors classification algorithm

---

**Input:**

$D$ : a set of training samples  $\{(\mathbf{x}_1, y_1), \dots, (\mathbf{x}_n, y_n)\}$

$k$ : the number of nearest neighbors

$d(\mathbf{x}, \mathbf{y})$ : a distance metric

$\mathbf{x}$ : a test sample

1: **for each** training sample  $(\mathbf{x}_i, y_i) \in D$  **do**

2:     Compute  $d(\mathbf{x}, \mathbf{x}_i)$ , the distance between  $\mathbf{x}$  and  $\mathbf{x}_i$

3: Let  $N \subseteq D$  be the set of training samples with the  $k$  smallest distances  $d(\mathbf{x}, \mathbf{x}_i)$

4: **return** the majority label of the samples in  $N$

---



# GIẢI THUẬT K-NEAREST NEIGHBOR (K-NN)

- Tính khoảng cách giữa 2 vector với Euclidean

We first compute the distance from  $(1, 0, 1)$  to each one of the given points.

$$d(\mathbf{x}, \mathbf{y}) = \sqrt{(x_1 - y_1)^2 + \dots + (x_n - y_n)^2} = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$

- Hoặc biểu diễn khác

$$d(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|$$

Some computations by computing the Euclidean



# GIẢI THUẬT K-NEAREST NEIGHBOR (K-NN)

- Ví dụ: Phân lớp vector  $i_9 = (1, 0, 1)$ ,  $k = 3$ , từ tập mẫu train

$i$	$x_1$	$x_2$	$x_3$	$y$
1	1	4	1	1
2	1	0	-2	0
3	0	0	1	0
4	-1	4	0	1
5	-1	-1	1	1
6	1	2	3	1
7	0	-4	0	0
8	1	0	-3	0



# GIẢI THUẬT WEIGHTED K-NEAREST NEIGHBOR (K-NN)

- Ví dụ: Phân lớp vector  $i_9 = (1, 0, 1)$ ,  $k = 3$ , từ tập mẫu train

$i$	$x_1$	$x_2$	$x_3$	$y$
1	1	4	1	1
2	1	0	-2	0
3	0	0	1	0
4	-1	4	0	1
5	-1	-1	1	1
6	1	2	3	1
7	0	-4	0	0
8	1	0	-3	0

B1. Tính khoảng cách từng mẫu  $i$  với  $i_9$

$$d(i_1, i_9) = \sqrt{(1 - 1)^2 + (4 - 0)^2 + (1 - 1)^2} = 4$$

$$d(i_2, i_9) = \sqrt{(1 - 1)^2 + (0 - 0)^2 + (-2 - 1)^2} = 3$$

$$d(i_3, i_9) = \sqrt{(0 - 1)^2 + (0 - 0)^2 + (1 - 1)^2} = 1$$

$$d(i_4, i_9) = \sqrt{(-1 - 1)^2 + (4 - 0)^2 + (0 - 1)^2} = 4.358$$

$$d(i_5, i_9) = \sqrt{(-1 - 1)^2 + (-1 - 0)^2 + (1 - 1)^2} = 2.236$$

$$d(i_6, i_9) = \sqrt{(1 - 1)^2 + (2 - 0)^2 + (3 - 1)^2} = 2.828$$

$$d(i_7, i_9) = \sqrt{(0 - 1)^2 + (-4 - 0)^2 + (0 - 1)^2} = 4.242$$

$$d(i_8, i_9) = \sqrt{(1 - 1)^2 + (0 - 0)^2 + (-3 - 1)^2} = 4$$



# GIẢI THUẬT WEIGHTED K-NEAREST NEIGHBOR (K-NN)

- Ví dụ: Phân lớp vector  $i_9 = (1, 0, 1)$ ,  $k = 3$ , từ tập mẫu train

$i$	$x_1$	$x_2$	$x_3$	$y$
1	1	4	1	1
2	1	0	-2	0
3	0	0	1	0
4	-1	4	0	1
5	-1	-1	1	1
6	1	2	3	1
7	0	-4	0	0
8	1	0	-3	0

B2. Chọn nhãn có số lượng nhiều nhất làm nhãn dự báo

$$d(i_3, i_9) = \sqrt{(0 - 1)^2 + (0 - 0)^2 + (1 - 1)^2} = 1$$

$$d(i_5, i_9) = \sqrt{(-1 - 1)^2 + (-1 - 0)^2 + (1 - 1)^2} = 2.236$$

$$d(i_6, i_9) = \sqrt{(1 - 1)^2 + (2 - 0)^2 + (3 - 1)^2} = 2.828$$

Mẫu	Khoản cách	Nhãn lớp
6	2.828	1
5	2.236	1
3	1	0

→  $i_9$  thuộc nhãn 1



# BÀI TẬP GIẢI THUẬT K-NEAREST NEIGHBOR (K-NN)

- Bài 1: dự báo kết quả (Result) của SV9 = (6.1, 40, 5), k = 3

No.	CGPA	Assessment	Project submitted	Result
1	9.2	85	8	Pass
2	8	80	7	Pass
3	8.5	81	8	Pass
4	6	45	4	Fail
5	6.5	50	4	Fail
6	8.2	72	7	Pass
7	5.8	38	5	Fail
8	8.9	91	9	Pass



# BÀI TẬP GIẢI THUẬT K-NEAREST NEIGHBOR (K-NN)

- Bài 2: dự báo Name (Result) của 2 loài thực vật có hoa sau, k = 3

SepalLength	SepalWidth	PetalLength	PetalWidth	Name
4.6	3.2	1.4	0.2	Iris-setosa
5.3	3.7	1.5	0.2	Iris-setosa
5	3.3	1.4	0.2	Iris-setosa
7	3.2	4.7	1.4	Iris-versicolor
6.4	3.2	4.5	1.5	Iris-versicolor
6.9	3.1	4.9	1.5	Iris-versicolor
6.3	3.3	6	2.5	Iris-virginica
5.8	2.7	5.1	1.9	Iris-virginica
7.1	3	5.9	2.1	Iris-virginica
5	2.4	3.5	1.1	?
4.1	3	2.1	1	?



# BÀI TẬP GIẢI THUẬT K-NEAREST NEIGHBOR (K-NN)

- Bài 3: dự báo giá nhà (House Price Index) cho người sau, k = 3
- (Tính trung bình cộng tập k mẫu)

Age	Loan	House Price Index
25	\$40,000	135
35	\$60,000	256
45	\$80,000	231
20	\$20,000	267
35	\$120,000	139
52	\$18,000	150
23	\$95,000	127
40	\$62,000	216
60	\$100,000	139
48	\$220,000	250
33	\$150,000	264
48	\$142,000	?



# BÀI TẬP

## GIẢI THUẬT K-NEAREST NEIGHBOR (K-NN)

- Bài 4: dự báo nhãn Lenses cho người 11-20, k = 3

Record ID	Age	Spectacle prescription	Astigmatic	Tear production Rate	Class label Lenses
1	Young	Myope	No	Reduced	Noncontact
2	Young	Myope	No	Normal	Soft contact
3	Young	Myope	Yes	Reduced	Noncontact
4	Young	Myope	Yes	Normal	Hard contact
5	Young	Hypermetrope	No	Reduced	Noncontact
6	Young	Hypermetrope	No	Normal	Soft contact
7	Young	Hypermetrope	Yes	Reduced	Noncontact
8	Young	Hypermetrope	Yes	Normal	Hard contact
9	Pre-presbyopic	Myope	No	Reduced	Noncontact
10	Pre-presbyopic	Myope	No	Normal	Soft contact
11	Pre-presbyopic	Myope	Yes	Reduced	
12	Pre-presbyopic	Myope	Yes	Normal	
13	Pre-presbyopic	Hypermetrope	No	Reduced	
14	Pre-presbyopic	Hypermetrope	No	Normal	
15	Pre-presbyopic	Hypermetrope	Yes	Reduced	
16	Pre-presbyopic	Hypermetrope	Yes	Normal	
17	Presbyopic	Myope	No	Reduced	
18	Presbyopic	Myope	No	Normal	
19	Presbyopic	Myope	Yes	Reduced	
20	Presbyopic	Myope	Yes	Normal	



# GIẢI THUẬT WEIGHTED K-NEAREST NEIGHBOR (W-K-NN)

- Là phương pháp mở rộng của k-NN.
- Tập k mẫu được gán trọng số cao hơn khi so sánh với các trường hợp xa với trường hợp đang test.
- B1. Tính khoảng cách mẫu test với tất cả mẫu train
- B2. Chọn top k mẫu gần mẫu test
- B3. Tính nghịch đảo khoảng cách tập k-mẫu
- B4. Tính tổng khoảng cách nghịch đảo trong tập k-mẫu
- B5. Tính trọng số (weight) khoảng cách nghịch đảo trong tập k-mẫu
- B6. Cộng trọng số theo nhãn lớp → Dự báo nhãn lớp



# GIẢI THUẬT WEIGHTED K-NEAREST NEIGHBOR (K-NN)

- Ví dụ: Phân lớp vector  $i_9 = (1, 0, 1)$ ,  $k = 3$ , từ tập mẫu train

$i$	$x_1$	$x_2$	$x_3$	$y$
1	1	4	1	1
2	1	0	-2	0
3	0	0	1	0
4	-1	4	0	1
5	-1	-1	1	1
6	1	2	3	1
7	0	-4	0	0
8	1	0	-3	0

B1. Tính khoảng cách từng mẫu  $i$  với  $i_9$

$$d(i_1, i_9) = \sqrt{(1 - 1)^2 + (4 - 0)^2 + (1 - 1)^2} = 4$$

$$d(i_2, i_9) = \sqrt{(1 - 1)^2 + (0 - 0)^2 + (-2 - 1)^2} = 3$$

$$d(i_3, i_9) = \sqrt{(0 - 1)^2 + (0 - 0)^2 + (1 - 1)^2} = 1$$

$$d(i_4, i_9) = \sqrt{(-1 - 1)^2 + (4 - 0)^2 + (0 - 1)^2} = 4.358$$

$$d(i_5, i_9) = \sqrt{(-1 - 1)^2 + (-1 - 0)^2 + (1 - 1)^2} = 2.236$$

$$d(i_6, i_9) = \sqrt{(1 - 1)^2 + (2 - 0)^2 + (3 - 1)^2} = 2.828$$

$$d(i_7, i_9) = \sqrt{(0 - 1)^2 + (-4 - 0)^2 + (0 - 1)^2} = 4.242$$

$$d(i_8, i_9) = \sqrt{(1 - 1)^2 + (0 - 0)^2 + (-3 - 1)^2} = 4$$



# GIẢI THUẬT WEIGHTED K-NEAREST NEIGHBOR (K-NN)

- Ví dụ: Phân lớp vector  $i_9 = (1, 0, 1)$ ,  $k = 3$ , từ tập mẫu train

$i$	$x_1$	$x_2$	$x_3$	$y$
1	1	4	1	1
2	1	0	-2	0
3	0	0	1	0
4	-1	4	0	1
5	-1	-1	1	1
6	1	2	3	1
7	0	-4	0	0
8	1	0	-3	0

B2. Chọn tập mẫu  $k=3$  gần vector  $i_9$

$$d(i_3, i_9) = \sqrt{(0 - 1)^2 + (0 - 0)^2 + (1 - 1)^2} = 1$$

$$d(i_5, i_9) = \sqrt{(-1 - 1)^2 + (-1 - 0)^2 + (1 - 1)^2} = 2.236$$

$$d(i_6, i_9) = \sqrt{(1 - 1)^2 + (2 - 0)^2 + (3 - 1)^2} = 2.828$$



# GIẢI THUẬT WEIGHTED K-NEAREST NEIGHBOR (K-NN)

- Ví dụ: Phân lớp vector  $i_9 = (1, 0, 1)$ ,  $k = 3$ , từ tập mẫu train

$i$	$x_1$	$x_2$	$x_3$	$y$
1	1	4	1	1
2	1	0	-2	0
3	0	0	1	0
4	-1	4	0	1
5	-1	-1	1	1
6	1	2	3	1
7	0	-4	0	0
8	1	0	-3	0

B3. Tính nghịch đảo khoảng cách k tập mẫu được chọn

Mẫu	Khoảng cách Euclidean	Khoảng cách nghịch đảo (1/Euclidean)	Nhãn Lớp
6	2.828	0.353	1
5	2.236	0.447	1
3	1	1	0



# GIẢI THUẬT WEIGHTED K-NEAREST NEIGHBOR (K-NN)

- Ví dụ: Phân lớp vector  $i_9 = (1, 0, 1)$ ,  $k = 3$ , từ tập mẫu train

$i$	$x_1$	$x_2$	$x_3$	$y$
1	1	4	1	1
2	1	0	-2	0
3	0	0	1	0
4	-1	4	0	1
5	-1	-1	1	1
6	1	2	3	1
7	0	-4	0	0
8	1	0	-3	0

B4. Tính tổng nghịch đảo khoảng cách k tập mẫu được chọn

Mẫu	Khoảng cách Euclidean	Khoảng cách nghịch đảo	Nhãn Lớp
6	2.828	0.353	1
5	2.236	0.447	1
3	1	1	0

$$\text{Sum} = 0.353 + 0.447 + 1 = 1.8$$



# GIẢI THUẬT WEIGHTED K-NEAREST NEIGHBOR (K-NN)

- Ví dụ: Phân lớp vector  $i_9 = (1, 0, 1)$ ,  $k = 3$ , từ tập mẫu train

$i$	$x_1$	$x_2$	$x_3$	$y$
1	1	4	1	1
2	1	0	-2	0
3	0	0	1	0
4	-1	4	0	1
5	-1	-1	1	1
6	1	2	3	1
7	0	-4	0	0
8	1	0	-3	0

B5. Tính trọng số nghịch đảo khoảng cách k tập mẫu được chọn

Mẫu	Khoảng cách Euclidean	Khoảng cách nghịch đảo	Trọng số (Inverse/Sum)	Nhân Lớp
6	2.828	0.353	0.196	1
5	2.236	0.447	0.248	1
3	1	1	0.555	0

$$\text{Sum} = 0.353 + 0.447 + 1 = 1.8$$



# GIẢI THUẬT WEIGHTED K-NEAREST NEIGHBOR (K-NN)

- Ví dụ: Phân lớp vector  $i_9 = (1, 0, 1)$ ,  $k = 3$ , từ tập mẫu train

$i$	$x_1$	$x_2$	$x_3$	$y$
1	1	4	1	1
2	1	0	-2	0
3	0	0	1	0
4	-1	4	0	1
5	-1	-1	1	1
6	1	2	3	1
7	0	-4	0	0
8	1	0	-3	0

B5. Cộng trọng số nghịch đảo theo nhãn lớp

Mẫu	Khoảng cách Euclidean	Khoảng cách nghịch đảo	Trọng số (Inverse/Sum)	Nhãn Lớp
6	2.828	0.353	0.196	1
5	2.236	0.447	0.248	1
3	1	1	0.555	0

Lớp 1:  $0.196 + 0.248 = 0.444$

Lớp 0: 0.555

Kết luận:  $i_9 (1, 0, 1) \rightarrow$  có nhãn 0



# BÀI TẬP GIẢI THUẬT WEIGHTED K-NEAREST NEIGHBOR (K-NN)

Bài 1:

Phân lớp vector a, b sau

a = (enough, moderate, very good, great)

b = (normal, high, very good, middling)

Clothes	Temperature	Humidity	Wind	Class
more	high	prodigious	no	N
more	high	prodigious	great	N
more	high	prodigious	middling	N
normal	high	prodigious	no	P
normal	high	prodigious	middling	P
enough	moderate	prodigious	no	N
enough	moderate	prodigious	middling	N
enough	high	very good	no	P
enough	high	very good	great	N
more	moderate	prodigious	no	N
more	moderate	prodigious	middling	N
enough	moderate	very good	no	N
enough	moderate	very good	middling	N
more	moderate	very good	middling	P
more	moderate	very good	great	P
normal	moderate	prodigious	great	P
normal	moderate	prodigious	middling	P
normal	high	very good	no	P



# BÀI TẬP GIẢI THUẬT WEIGHTED K-NEAREST NEIGHBOR (K-NN)

Bài 2:

Phân lớp vector a, b sau

a = (Middle\_aged, High, Yes, Fair)

b = (Senior, Medium, No, Excellent)

RID	Age	Income	Student	Credit_rating	Class: buys computer
1	Youth	High	No	Fair	No
2	Youth	High	No	Excellent	No
3	middle_aged	High	No	Fair	Yes
4	Senior	Medium	No	Fair	Yes
5	Senior	Low	Yes	Fair	Yes
6	Senior	Low	Yes	Excellent	No
7	middle_aged	Low	Yes	Excellent	Yes
8	Youth	Medium	No	Fair	No
9	Youth	Low	Yes	Fair	Yes
10	Senior	Medium	Yes	Fair	Yes
11	Youth	Medium	Yes	Excellent	Yes
12	middle_aged	Medium	No	Excellent	Yes



# BÀI TẬP GIẢI THUẬT WEIGHTED K-NEAREST NEIGHBOR (K-NN)

Bài 3:

Phân lớp vector a, b sau, k = 4

a = (34, F, High, Normal)

b = (43, M, Low, High)

	Age	Sex	BP	Cholesterol	Drug
0	23	F	HIGH	HIGH	drugY
1	47	M	LOW	HIGH	drugC
2	47	M	LOW	HIGH	drugC
3	28	F	NORMAL	HIGH	drugX
4	61	F	LOW	HIGH	drugY
5	22	F	NORMAL	HIGH	drugX
6	49	F	NORMAL	HIGH	drugY
7	41	M	LOW	HIGH	drugC
8	60	M	NORMAL	HIGH	drugY
9	43	M	LOW	NORMAL	drugY
10	47	F	LOW	HIGH	drugC



# GIẢI THUẬT NEAREST CENTROID K-NEAREST NEIGHBOR (NC-K-NN)

- Là phương pháp mở rộng của k-NN.
- Ý tưởng: Phân lớp mẫu test có giá trị trung bình gần với nó.
- Giải thuật

**Inputs:** Training dataset  $T$ , Distance metric  $d$ , Test instance  $t$

**Output:** Predicted class or category

1. Compute the mean/centroid of each class.
2. Compute the distance between the test instance and mean/centroid of each class (Euclidean Distance).
3. Predict the class by choosing the class with the smaller distance.



# GIẢI THUẬT NEAREST CENTROID K-NEAREST NEIGHBOR (NC-K-NN)

- Ví dụ: Dự báo lớp cho mẫu  $x_i = (6, 5)$
- Bước 1: Tính trung bình mỗi lớp

$$\text{Centroid of class 'A'} = (3 + 5 + 4, 1 + 2 + 3)/3 = (12, 6)/3 = (4, 2)$$

$$\text{Centroid of class 'B'} = (7 + 6 + 8, 6 + 7 + 5)/3 = (21, 18)/3 = (7, 6)$$

X	Y	Class
3	1	A
5	2	A
4	3	A
7	6	B
6	7	B
8	5	B

# GIẢI THUẬT NEAREST CENTROID K-NEAREST NEIGHBOR (NC-K-NN)

- Ví dụ: Dự báo lớp cho mẫu  $x_i = (6, 5)$
- Bước 2: Tính khoảng cách trung bình mỗi lớp với mẫu test  $x_i$

$$\text{Euc\_Dist}[(6, 5) ; (4, 2)] = \sqrt{(6-4)^2 + (5-2)^2} = \sqrt{13} = 3.6$$

$$\text{Euc\_Dist}[(6, 5) ; (7, 6)] = \sqrt{(6-7)^2 + (5-6)^2} = \sqrt{2} = 1.414$$

- $\rightarrow x_i$  lớp B

X	Y	Class
3	1	A
5	2	A
4	3	A
7	6	B
6	7	B
8	5	B





# BÀI TẬP GIẢI THUẬT NEAREST CENTROID K-NEAREST NEIGHBOR (K-NN)

- Bài 1: dự báo kết quả (Result) của SV9 = (6.1, 40, 5)

No.	CGPA	Assessment	Project submitted	Result
1	9.2	85	8	Pass
2	8	80	7	Pass
3	8.5	81	8	Pass
4	6	45	4	Fail
5	6.5	50	4	Fail
6	8.2	72	7	Pass
7	5.8	38	5	Fail
8	8.9	91	9	Pass



# BÀI TẬP GIẢI THUẬT NEAREST CENTROID K-NEAREST NEIGHBOR (K-NN)

- Bài 2: dự báo Name (Result) của 2 loài thực vật có hoa sau

SepalLength	SepalWidth	PetalLength	PetalWidth	Name
4.6	3.2	1.4	0.2	Iris-setosa
5.3	3.7	1.5	0.2	Iris-setosa
5	3.3	1.4	0.2	Iris-setosa
7	3.2	4.7	1.4	Iris-versicolor
6.4	3.2	4.5	1.5	Iris-versicolor
6.9	3.1	4.9	1.5	Iris-versicolor
6.3	3.3	6	2.5	Iris-virginica
5.8	2.7	5.1	1.9	Iris-virginica
7.1	3	5.9	2.1	Iris-virginica
5	2.4	3.5	1.1	?
4.1	3	2.1	1	?



# BÀI TẬP TỔNG HỢP

Bài 1: phân lớp mẫu test  $x_{11} = (7.8, 4)$ , k = 3 (Dùng 3 giải thuật)

S.No.	GPA	No. of Projects Done	Award
1.	9.5	5	Yes
2.	8.0	4	Yes
3.	7.2	1	No
4.	6.5	5	Yes
5.	9.5	4	Yes
6.	3.2	1	No
7.	6.6	1	No
8.	5.4	1	No
9.	8.9	3	Yes
10.	7.2	4	Yes

# BÀI TẬP TỔNG HỢP

Bài 2: phân lớp mẫu test  $x_{11} = (\text{Yes}, \text{Yes}, \text{Yes}, \text{Yes}, \text{No}, \text{No}, \text{No}, \text{No}, \text{No})$

S.No.	Fever	Dry cough	Tiredness	Sore Throat	Diarrhea	Headache	Loss of Taste or Smell	Shortness of Breath	Chest Pain	Result
1.	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Positive
2.	Yes	No	Yes	No	No	Yes	No	No	No	Negative
3.	No	No	No	No	No	No	No	No	No	Negative
4.	Yes	Yes	Yes	No	No	No	No	No	Yes	Negative
5.	Yes	Yes	Yes	Yes	No	No	Yes	Yes	Yes	Positive
6.	Yes	Yes	Yes	Yes	No	Yes	No	No	No	Positive
7.	Yes	Yes	Yes	Yes	No	No	Yes	Yes	No	Positive
8.	Yes	Yes	Yes	Yes	No	No	No	No	No	Positive
9.	Yes	Yes	Yes	Yes	No	No	No	No	No	Positive
10.	No	No	No	No	No	No	No	No	No	Negative





**Q&A**

