

[Trang chủ](#) / [Khóa học](#) / [Học Kỳ I năm học 2021-2022 \(Semester 1 - Academic year 2021-2022\)](#)

/ [Đại Học Chính Quy \(Bachelor program \(Full-time study\)\)](#)

/ [Khoa Khoa học và Kỹ thuật Máy tính \(Faculty of Computer Science and Engineering.\)](#) / [Hệ Thống Thông Tin](#)

/ [CO3029_20/12/2021_07g00_DH_HK211](#) / [General](#) / [Thi cuối kỳ - KH211](#)

Thời gian còn lại 0:00:30

Câu hỏi **1**

Câu trả lời đã được lưu

Chấm điểm của 0,33

Câu hỏi 1 đến 6 xét bài toán phân cụm sau đây. Giả sử ta cần phân cụm 8 điểm dữ liệu $A1 = (2,10)$, $A2 = (2,5)$, $A3 = (8,4)$, $A4 = (5,8)$, $A5 = (7,5)$, $A6 = (6,4)$, $A7 = (1,2)$ và $A8 = (4,9)$ thành 3 cụm có tâm lần lượt là $C1$, $C2$ và $C3$, dùng thuật toán k-means với độ đo là khoảng cách Euclidean.

Khởi tạo các tâm lần lượt là $C1 = A1$, $C2 = A4$ và $C3 = A7$. Tại lần lặp đầu tiên, các cụm tìm được là

- ☐ a. $\{A1, A3\}, \{A4, A5, A6, A8\}, \{A2, A7\}$
- ☐ b. $\{A1, A4, A8\}, \{A3, A5, A6\}, \{A2, A7\}$
- ☒ c. $\{A1\}, \{A3, A4, A5, A6, A8\}, \{A2, A7\}$
- ☐ d. $\{A1, A8\}, \{A3, A4, A5, A6\}, \{A2, A7\}$

[Clear my choice](#)

Câu hỏi **2**

Câu trả lời đã được lưu

Chấm điểm của 0,33

Sau lần lặp đầu tiên, các tâm mới của cụm được cập nhật tương ứng là

- ☐ a. $C1 = (2,10)$, $C2 = (6,6)$, $C3 = (3.5,1.5)$
- ☐ b. $C1 = (6,6)$, $C2 = (2,10)$, $C3 = (1.5,3.5)$
- ☒ c. $C1 = (2,10)$, $C2 = (6,6)$, $C3 = (1.5,3.5)$
- ☐ d. $C1 = (10,2)$, $C2 = (6,6)$, $C3 = (1.5,3.5)$

[Clear my choice](#)Câu hỏi **3**

Câu trả lời đã được lưu

Chấm điểm của 0,33

Tại lần lặp thứ 2, các cụm tìm được là

- ☐ a. $\{A1.A3\}$, $\{A4, A5, A6, A8\}$, $\{A2, A7\}$
- ☒ b. $\{A1, A8\}$, $\{A3, A4, A5, A6\}$, $\{A2, A7\}$
- ☐ c. $\{A1, A4, A8\}$, $\{A3, A5, A6\}$, $\{A2, A7\}$
- ☐ d. $\{A1\}$, $\{A3, A4, A5, A6, A8\}$, $\{A2, A7\}$

[Clear my choice](#)

Câu hỏi **4**

Câu trả lời đã được lưu

Chấm điểm của 0,33

Sau lần lặp thứ 2, các tâm mới của cụm được cập nhật tương ứng là

- ☐ a. $C1 = (3, 9.5), C2 = (6.5, 5.25), C3 = (3.5, 1.5)$
- ☐ b. $C1 = (9.5, 3), C2 = (5.5, 5.25), C3 = (1.5, 3.5)$
- ☒ c. $C1 = (3, 9.5), C2 = (6.5, 5.25), C3 = (1.5, 3.5)$
- ☐ d. $C1 = (9.5, 3), C2 = (6.5, 5.25), C3 = (1.5, 3.5)$

[Clear my choice](#)Câu hỏi **5**

Câu trả lời đã được lưu

Chấm điểm của 0,33

Kết quả phân cụm cuối cùng là

- ☐ a. $\{A4, A8\}, \{A3, A5, A6\}, \{A1, A2, A7\}$
- ☐ b. $\{A1\}, \{A3, A4, A5, A6, A8\}, \{A2, A7\}$
- ☒ c. $\{A1, A4, A8\}, \{A3, A5, A6\}, \{A2, A7\}$
- ☐ d. $\{A1, A8\}, \{A3, A4, A5, A6\}, \{A2, A7\}$

[Clear my choice](#)

Câu hỏi **6**

Câu trả lời đã được lưu

Chấm điểm của 0,33

Tâm của các cụm tìm được cuối cùng là

- ☐ a. $C1 = (9.5, 3)$, $C2 = (5.5, 5.25)$, $C3 = (1.5, 3.5)$
- ☐ b. $C1 = (9.5, 3)$, $C2 = (6.5, 5.25)$, $C3 = (1.5, 3.5)$
- ☐ c. $C1 = (3.66, 9)$, $C2 = (4.33, 7)$, $C3 = (1.5, 3.5)$
- ☒ d. $C1 = (3.66, 9)$, $C2 = (7, 4.33)$, $C3 = (1.5, 3.5)$

[Clear my choice](#)

Câu hỏi **7**

Câu trả lời đã được lưu

Chấm điểm của 0,33

Các câu hỏi 7 đến 11 xét danh sách giao dịch dưới đây:

- (1) trứng, thịt, khử khuẩn, khẩu trang
- (2) sữa, khẩu trang, trứng, mỳ gói, thịt
- (3) khẩu trang, sữa, trứng
- (4) mỳ gói, bia, sữa
- (5) khử khuẩn, mỳ gói, bia, sữa
- (6) thịt
- (7) sữa, bia, mỳ gói, khẩu trang, thịt
- (8) trứng, thịt, bia
- (9) sữa, mỳ gói
- (10) khẩu trang, sữa, mỳ gói

Danh sách có

- ☐ a. 7 giao dịch
- ☒ b. 10 giao dịch
- ☐ c. 9 giao dịch
- ☐ d. 8 giao dịch

[Clear my choice](#)

Câu hỏi 8

Câu trả lời đã được lưu

Chấm điểm của 0,33

Với $support = 0.4$, danh sách các mẫu (itemsets) xuất hiện thường xuyên là

- ☐ a. {mỳ gói, khẩu trang}, {khẩu trang, bia, thịt}, {sữa, mỳ gói, trứng}
- ☐ b. {trứng, mỳ gói}, {bia, thịt}, {sữa, mỳ gói, trứng}
- ☐ c. {trứng, mỳ gói}, {bia, khẩu trang}, {mỳ gói, khẩu trang}
- ☒ d. {trứng}, {thịt}, {khẩu trang}, {sữa}, {mỳ gói}, {bia}, {khẩu trang, sữa}, {sữa, mỳ gói}

[Clear my choice](#)

Câu hỏi 9

Câu trả lời đã được lưu

Chấm điểm của 0,33

Nếu giảm giá trị của $support$ xuống, thì

- ☐ a. số mẫu xuất hiện thường xuyên vẫn luôn giữ nguyên
- ☐ b. một số mẫu sẽ được đưa ra khỏi tập xuất hiện thường xuyên hiện tại
- ☐ c. các phương án đều đúng
- ☒ d. một số mẫu có thể được thêm vào tập xuất hiện thường xuyên hiện tại

[Clear my choice](#)

Câu hỏi **10**

Câu trả lời đã được lưu

Chấm điểm của 0,33

Các luật kết hợp với $support = 0.4$ và $confidence = 0.7$ gồm

- ☐ a. $\{sữa\} \rightarrow \{bia\}, \{bia\} \rightarrow \{khẩu\ trang\}$
- ☐ b. $\{bia\} \rightarrow \{khẩu\ trang\}, \{khẩu\ trang\} \rightarrow \{bia\}$
- ☒ c. $\{khẩu\ trang\} \rightarrow \{sữa\}, \{sữa\} \rightarrow \{mỳ\ gói\}, \{mỳ\ gói\} \rightarrow \{sữa\}$
- ☐ d. $\{khẩu\ trang\} \rightarrow \{mỳ\ gói\}, \{mỳ\ gói\} \rightarrow \{khẩu\ trang\}$

[Clear my choice](#)Câu hỏi **11**

Câu trả lời đã được lưu

Chấm điểm của 0,33

Kết quả khai phá luật kết hợp thu được cho thấy

- ☐ a. bia và khẩu trang thường sẽ được mua cùng nhau
- ☐ b. sữa và bia thường sẽ được mua cùng nhau
- ☐ c. mỳ gói và khẩu trang thường sẽ được mua cùng nhau
- ☒ d. sữa và mỳ gói thường sẽ được mua cùng nhau

[Clear my choice](#)

Câu hỏi **12**

Câu trả lời đã được lưu

Chấm điểm của 0,33

Kỹ thuật nào dưới đây thích hợp nhất khi áp dụng để xác định một bài viết (trên mạng xã hội) được thích hay không?

- ☐ a. Phân cụm
- ☐ b. Khai phá luật kết hợp
- ☐ c. Hồi quy
- ☒ d. Phân lớp

[Clear my choice](#)Câu hỏi **13**

Câu trả lời đã được lưu

Chấm điểm của 0,33

Bộ phân loại Naive Bayes có thời gian huấn luyện và dự đoán/kiểm tra nhanh. Có được điều này là do

- ☐ a. Các thành phần phụ thuộc lẫn nhau
- ☐ b. Dữ liệu huấn luyện thường có kích thước nhỏ
- ☒ c. Giả thiết về tính độc lập giữa các thành phần
- ☐ d. Bộ phân lớp này sử dụng mô hình tiền huấn luyện (pre-trained)

[Clear my choice](#)Câu hỏi **14**

Câu trả lời đã được lưu

Chấm điểm của 0,33

Khi áp dụng k-means để phân đoạn ảnh (image segmenting) mỗi điểm ảnh sẽ là một điểm dữ liệu được mô tả

- ☒ a. bởi một vector ba chiều chứa các giá trị màu R, G và B
- ☐ b. bởi căn bậc 2 của độ lớn của vector ba chiều chứa các giá trị màu R, G và B
- ☐ c. bởi một giá trị màu trong khoảng 0 đến 255
- ☐ d. bởi độ lớn của vector ba chiều chứa các giá trị màu R, G và B

[Clear my choice](#)

Câu hỏi **15**

Câu trả lời đã được lưu

Chấm điểm của 0,33

Việc phân cụm không yêu cầu

- ☐ a. Dữ liệu không gán nhãn
- ☐ b. Dữ liệu văn bản
- ☐ c. Dữ liệu số
- ☒ d. Dữ liệu được gán nhãn

[Clear my choice](#)Câu hỏi **16**

Câu trả lời đã được lưu

Chấm điểm của 0,33

Hàm kích hoạt sigmoid dùng cho bài toán phân lớp nhị phân có tổng các giá trị xác suất luôn bằng 1 ?

Chọn một:

- ☒ Đúng
- ☐ Sai

Câu hỏi **17**

Câu trả lời đã được lưu

Chấm điểm của 0,33

Lược đồ dendrogram biểu diễn một

- ☐ a. Tất cả đáp án đều đúng
- ☐ b. Cấu trúc đồ thị
- ☐ c. Cấu trúc sơ đồ
- ☒ d. Cấu trúc phân cấp

[Clear my choice](#)

Câu hỏi 18

Câu trả lời đã được lưu

Chấm điểm của 0,33

Hàm nào dưới đây có thể dùng làm hàm kích hoạt trong các mô hình mạng nơ-ron

☐ a.

$$f(x) = x$$

☐ b.

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

☐ c.

$$f(x) = \begin{cases} 0 & \text{for } x < 0 \\ 1 & \text{for } x \geq 0 \end{cases}$$

☒ d. Tất cả đều có thể sử dụng

[Clear my choice](#)

Câu hỏi 19

Câu trả lời đã được lưu

Chấm điểm của 0,33

Với một tập dữ liệu đầu vào cho trước, hai lần chạy của một thuật toán k-means với tập dữ liệu này sẽ luôn cho kết quả

☐ a. Giống nhau

☐ b. Khác nhau

☒ c. Có thể giống, có thể khác

☐ d. Không có đáp án nào đúng

[Clear my choice](#)

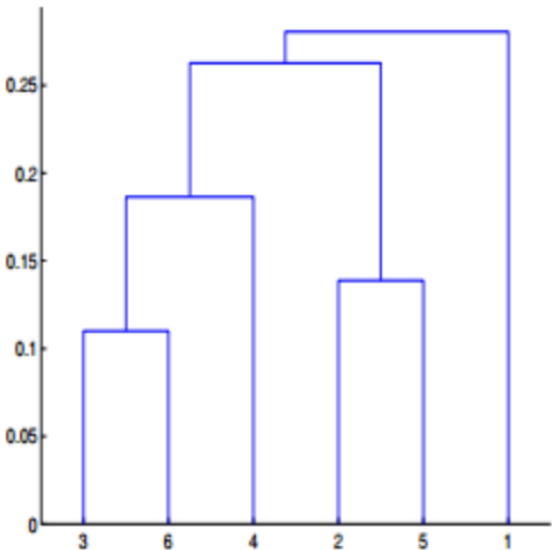
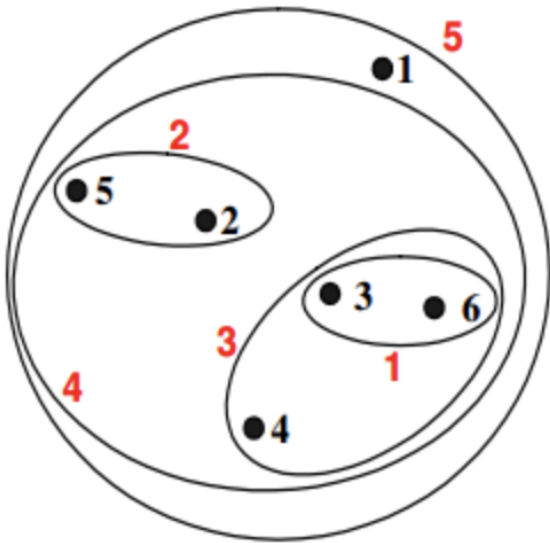
Câu hỏi 20

Câu trả lời đã được lưu

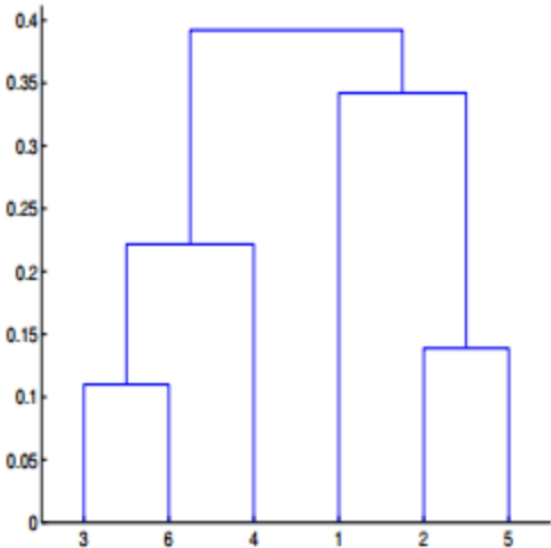
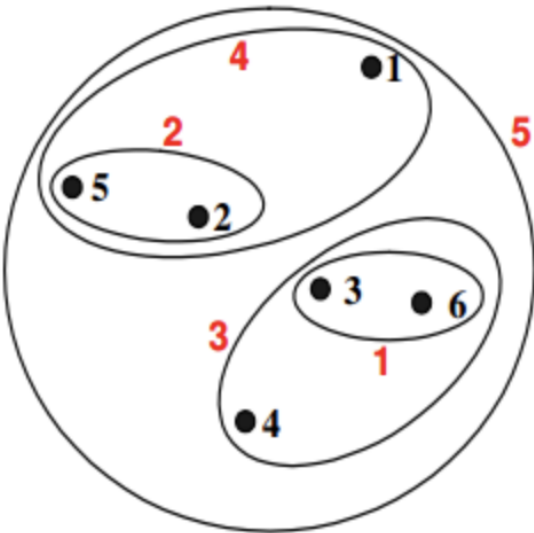
Chấm điểm của 0,41

Cho 6 điểm dữ liệu $p_1 = (0.4005, 0.5306)$, $p_2 = (0.2148, 0.3854)$, $p_3 = (0.3457, 0.3156)$, $p_4 = (0.2652, 0.1875)$, $p_5 = (0.0789, 0.4139)$ và $p_6 = (0.4548, 0.3022)$. Sử dụng khoảng cách single-link, hình ảnh biểu diễn cụm và cấu trúc phân cấp cụm của tập dữ liệu này là

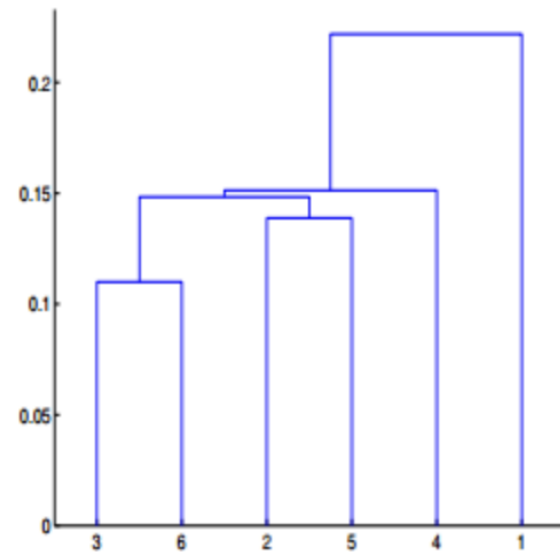
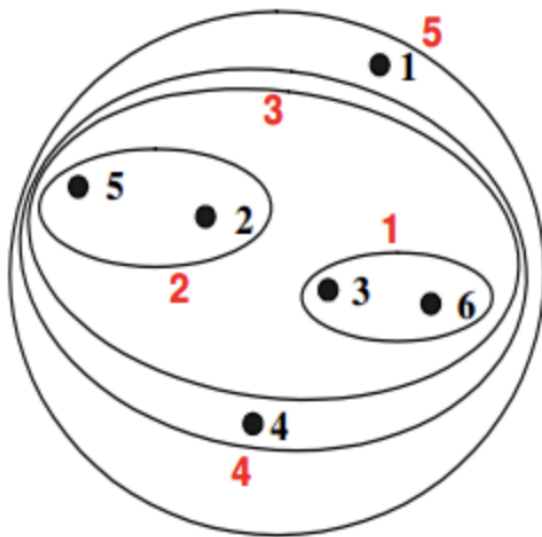
☐ a.



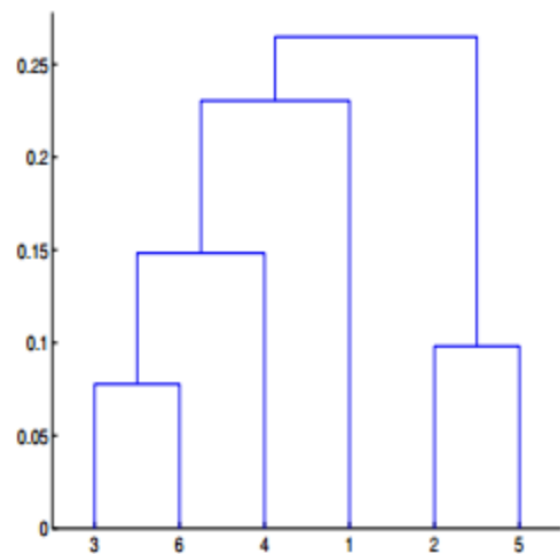
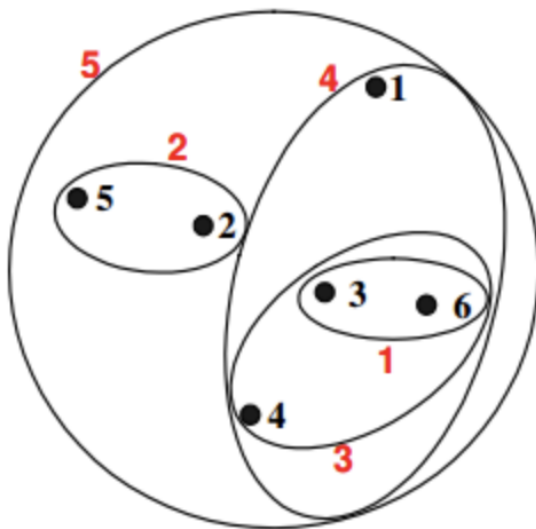
☐ b.



☒ c.



☐ d.



[Clear my choice](#)

Câu hỏi 21

Câu trả lời đã được lưu

Chấm điểm của 0,33

Nếu giá trị support của itemset $\{a, b\}$ là 10, thì giá trị support của $\{a, b, c\}$ có thể là

- ☐ a. 12
- ☐ b. 11
- ☐ c. 10
- ☒ d. 9

[Clear my choice](#)

Câu hỏi **22**

Câu trả lời đã được lưu

Chấm điểm của 0,33

Kỹ thuật khai phá nào dưới đây là phù hợp nhất cho bài toán: xác định sự đánh giá (rate) của một địa điểm du lịch cho trước?

- ☐ a. Dự đoán
- ☒ b. Phân lớp
- ☐ c. Phân cụm
- ☐ d. Khai phá luật kết hợp

[Clear my choice](#)Câu hỏi **23**

Câu trả lời đã được lưu

Chấm điểm của 0,33

Cho bảng giao dịch như dưới đây:

1. Kẹo bì, hạt, viết
2. Kẹo bì, cafe, viết, hạt
3. Kẹo bì, viết, trứng
4. Kẹo bì, hạt, trứng, sữa
5. Hạt, cafe, viết, trứng, sữa

Cho biết giá trị minsup là 50%, có bao nhiêu 3-itemsets xuất hiện thường xuyên?

- ☐ a. 3
- ☐ b. 2
- ☒ c. 0
- ☐ d. 1

[Clear my choice](#)

Câu hỏi **24**

Câu trả lời đã được lưu

Chấm điểm của 0,33

Bài toán nào dưới là ứng dụng trực tiếp của khai phá luật kết hợp?

- ☒ a. Phân tích giỏ hàng
- ☐ b. Phân tích mạng xã hội
- ☐ c. Phát hiện xâm nhập
- ☐ d. Phát hiện ngoại lai

[Clear my choice](#)Câu hỏi **25**

Câu trả lời đã được lưu

Chấm điểm của 0,33

Đại lượng $\text{Lift}(A \rightarrow B)$ được định nghĩa là

- ☐ a. $P(A|B)/P(B)$
- ☐ b. $P(A|B)/P(A)$
- ☐ c. $P(B|A)/P(A)$
- ☒ d. $P(B|A)/P(B)$

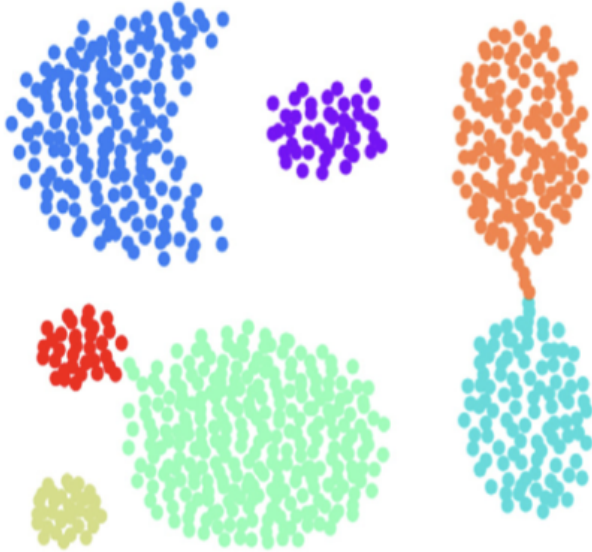
[Clear my choice](#)

Câu hỏi **26**

Câu trả lời đã được lưu

Chấm điểm của 0,33

Một tập dữ liệu được kỳ vọng sẽ được phân thành 7 cụm với các màu khác nhau như dưới đây.



Nếu áp dụng thuật toán DBSCAN với các tham số epsilon và minPts được lấy theo kinh nghiệm thông thường, số cụm nhận được sẽ thường là

- ☐ a. 6
- ☐ b. 8
- ☐ c. 5
- ☒ d. 7

[Clear my choice](#)Câu hỏi **27**

Câu trả lời đã được lưu

Chấm điểm của 0,33

Nguyên lý chính của giải thuật Apriori gồm

- ☐ a. (b) Giá trị support của một itemset không bao giờ vượt quá giá trị support của các tập con của nó
- ☐ b. (a) Nếu một Itemset xuất hiện thường xuyên thì tất cả tập con của Itemset này cũng xuất hiện thường xuyên
- ☒ c. Cả (a) và (b)

[Clear my choice](#)

Câu hỏi **28**

Câu trả lời đã được lưu

Chấm điểm của 0,33

Thách thức lớn nhất trong thuật toán Apriori nằm ở

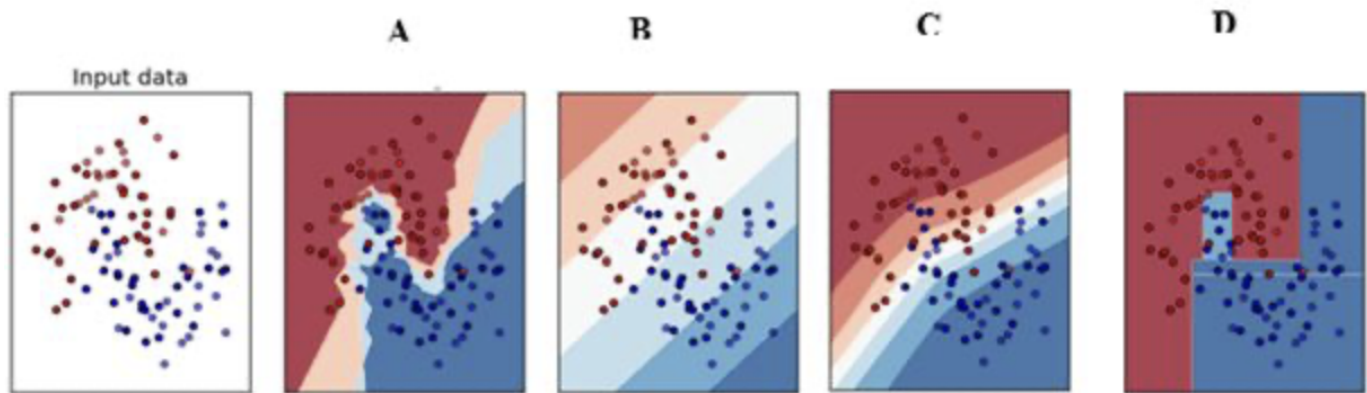
- ☐ a. Số lần lặp (iteration)
- ☐ b. Tất cả yếu tố này
- ☐ c. Việc cắt tỉa
- ☒ d. Việc sinh các ứng viên

[Clear my choice](#)Câu hỏi **29**

Câu trả lời đã được lưu

Chấm điểm của 0,33

Cho một tập dữ liệu vào và các hình ảnh kết quả phân lớp như hình dưới đây.



Hình ảnh nào trong số A, B, C, và D là thể hiện biên quyết định phân lớp bởi một mạng nơ-ron?

- ☐ a. C
- ☒ b. Tất cả A, B, C và D
- ☐ c. A
- ☐ d. D
- ☐ e. B

[Clear my choice](#)

Câu hỏi **30**

Câu trả lời đã được lưu

Chấm điểm của 0,33

Điều gì xảy ra nếu giá trị tham số epsilon trong thuật toán DBSCAN quá bé?

- ☐ a. Số lượng điểm nhiễu/ngoại lai sẽ tăng
- ☐ b. Số lượng cụm sẽ giảm
- ☐ c. Thuật toán không bị tác động gì cả
- ☒ d. Mật độ các cụm thu được sẽ dày đặc hơn

[Clear my choice](#)