

Time left

Question **27**

Not yet  
answered

Marked out of  
1.00

Hold-out có phải là một phương pháp để tiền xử lý và hiểu dữ liệu?

Select one:

- ☐ a. Không, nó là một phương pháp để huấn luyện một mô hình từ một tập dữ liệu cho trước
- ☐ b. Đúng, tất nhiên rồi
- ☒ c. Không, nó là một chiến lược để đánh giá một mô hình

[Previous page](#)

[Next page](#)

## on Học máy và khai phá dữ liệu

Time left 0:51:4

Question **35**

Not yet  
answered

Marked out of  
1.00

Cây quyết định ...

Select one:

- ☐ a. không phải là phương pháp học có giám sát và cũng không phải là phương pháp học không giám sát
- ☒ b. là một phương pháp học có giám sát
- ☐ c. là một phương pháp học không giám sát
- ☐ d. là một phương pháp học có giám sát và cũng là một phương pháp học không giám sát

[Previous page](#)

[Next page](#)



Question **34**

Not yet  
answered

Marked out of  
1.00

Một nút thắt cổ chai nghiêm trọng của thuật toán Apriori là

Select one:

- ☐ a. sinh ứng viên
- ☒ b. tìm các tập mục thường xuyên
- ☐ c. số vòng lặp
- ☐ d. cắt tỉa

[Previous page](#)

[Next page](#)



# n Học máy và khai phá dữ liệu

Time left 0

Question **29**

Not yet  
answered

Marked out of  
1.00

Phân loại Bayes là

Select one:

- ☐ a. Một lớp các thuật toán mà sử dụng information gain
- ☒ b. Một lớp các thuật toán mà sử dụng tính toán về tần suất xuất hiện của các giá trị thuộc tính trong tập huấn luyện
- ☐ c. Một lớp các thuật toán mà sử dụng các tính toán về xác suất hậu nghiệm

[Previous page](#)

[Next](#)



# h Học máy và khai phá dữ liệu

Time left 0:52

Question **32**

Not yet  
answered

Marked out of  
1.00

Trong mạng nơ-ron, đâu là lợi ích của hàm kích hoạt ReLU so với hàm kích hoạt Sigmoid?

Select one:

- ☒ a. ReLUs chấp nhận tính toán lan truyền ngược gradient nhanh hơn
- ☐ b. ReLUs có thể sử dụng cho tầng đầu ra, trong khi hàm sigmoid thì không thể
- ☐ c. ReLUs chấp nhận mô hình học một đường quyết định phi tuyến

Previous page

Next

# Machine Learning and Data Mining

Time left 0:54

Question **24**

Not yet  
answered

Marked out of  
1.00

Thuật toán K-means cho phân cụm đảm bảo hội tụ tới một điểm cực trị địa phương hoặc điểm dừng?

Select one:

☒ True

☐ False

[Previous page](#)

[Next](#)



Question **22**

Not yet  
answered

Marked out of  
1.00

Số vòng lặp của thuật toán Apriori

Select one:

- ☐ a. giảm khi kích thước tập mục thường xuyên lớn nhất tăng kích thước
- ☐ b. tăng theo kích thước của dữ liệu
- ☒ c. tăng theo kích thước của tập mục thường xuyên lớn nhất

[Previous page](#)

Time left 0:52:00

on 33

t  
red  
d out of

Phát biểu nào sau đây là đúng về giải thuật học láng giềng gần nhất?

Select one:

- ☐ a. Khi số lượng các thuộc tính kiểu số là ít, giải thuật học láng giềng gần nhất đạt độ chính xác cao hơn nếu các thuộc tính kiểu số đó có cùng khoảng giá trị.
- ☐ b. Khi số lượng các thuộc tính kiểu số là nhiều, giải thuật học láng giềng gần nhất đạt độ chính xác cao hơn nếu các thuộc tính kiểu số đó có cùng khoảng giá trị.
- ☒ c. Bất kể số lượng các thuộc tính kiểu số, giải thuật học láng giềng gần nhất đạt độ chính xác cao hơn nếu các thuộc tính kiểu số đó có cùng khoảng giá trị.

Next page

ious page



Question **30**

Not yet  
answered

Marked out of  
1.00

Lựa chọn nào sau đây không là một bài toán của Khai phá dữ liệu?

Select one:

- ☐ a. Khám phá dữ liệu (Data exploration)
- ☐ b. Phân cụm (Cluster analysis)
- ☒ c. Biến đổi dữ liệu (Data Transformation)
- ☐ d. Trích rút tri thức (Knowledge extraction)

[Previous page](#)

Question 10

Answer saved

Marked out of  
1.00

Học máy (Machine Learning) cung cấp các phương pháp để phân tích dữ liệu, tạo các phân đoán cho các quan sát trong tương lai

Select one:

- ☐ a. Đúng
- ☐ b. Sai, nó cung cấp các phương pháp tăng tốc tính toán của máy tính
- ☒ c. Đúng, nó còn cung cấp các phương pháp tăng tốc tính toán của máy tính

[Clear my choice](#)

[Previous page](#)

[Next page](#)



Với một cây quyết định học được, thì nhân phân loại đối với 1 ví dụ mới được xác định bởi việc duyệt (đi theo) ...

Select one:

- ☐ a. tất cả các đường đi trong cây
- ☒ b. một đường đi duy nhất trong cây
- ☐ c. đường đi có độ sâu lớn nhất trong cây
- ☐ d. đường đi có độ sâu nhỏ nhất trong cây

Clear my choice

Time left 0:07:59

Question 23

Answer saved

Marked out of  
1.00

Một biến quan sát được có thể biểu diễn

Select one:

- ☐ a. bất kể thứ gì
- ☐ b. một sự kiện mà có thể quan sát được hoặc không
- ☐ c. một sự kiện mà không thể quan sát được
- ☒ d. một sự kiện mà có thể quan sát được trong thực tế

[Clear my choice](#)

[Previous page](#)

[Next page](#)



Time left 0:07:59

Question 23

Answer saved

Marked out of  
1.00

Một biến quan sát được có thể biểu diễn

Select one:

- ☐ a. bất kể thứ gì
- ☐ b. một sự kiện mà có thể quan sát được hoặc không
- ☐ c. một sự kiện mà không thể quan sát được
- ☒ d. một sự kiện mà có thể quan sát được trong thực tế

[Clear my choice](#)

[Previous page](#)

[Next page](#)

Time left 0:07:23

Question 31

Answer saved

Marked out of  
1.00

Trong hard-margin SVM cho dữ liệu phân tách tuyến tính, các vector hỗ trợ nằm

Select one:

- ☐ a. nằm ở vùng giữa đường phân tách âm hoặc đường phân tách dương
- ☐ b. trên đường phân tách âm
- ☐ c. trên đường phân tách dương
- ☒ d. trên đường phân tách âm hoặc đường phân tách dương

Clear my choice

Previous page

Next page



Time left 0:07:23

Question 31

Answer saved

Marked out of 1.00

Trong hard-margin SVM cho dữ liệu phân tách tuyến tính, các vector hỗ trợ nằm

Select one:

- ☐ a. nằm ở vùng giữa đường phân tách âm hoặc đường phân tách dương
- ☐ b. trên đường phân tách âm
- ☐ c. trên đường phân tách dương
- ☒ d. trên đường phân tách âm hoặc đường phân tách dương

Clear my choice

Previous page

Next page

## Question 2

Answer saved

Marked out of  
1.00

Xét tập huấn luyện như sau ghi lại các trường hợp khách hàng có hay không mua máy tính. Mỗi ví dụ huấn luyện được biểu diễn bởi 3 thuộc tính:

- *Age*: mức độ tuổi của khách hàng.
- *Income*: mức độ thu nhập của khách hàng.
- *Buy a computer*: khách hàng mua hay không mua máy tính.

<i>Age</i>	<i>Income</i>	<i>Buy a computer</i>
Young	Low	No
Young	High	Yes
Medium	Low	Yes
Old	Low	No
Old	High	Yes
Old	High	Yes
Medium	High	Yes
Young	Low	No
Young	High	Yes
Old	Low	No

Áp dụng giải thuật học cây quyết định ID3, hãy xác định thuộc tính kiểm tra cho nút gốc của cây quyết định học được?

Select one:

- ☒ a. Age
- ☐ b. Income

Chọn một câu trả lời

# ây và khai phá dữ liệu

Time left 0:

Những phát biểu nào sau đây SAI về thuật toán K-means

Select one or more:

- ☐ a. K-means không thể bị quá khớp (overfitting) do không có hàm lỗi nào được sử dụng cho huấn luyện
- ☐ b. Hiệu năng của thuật toán K-means có thể thay đổi khi chạy các lần khác nhau với các khởi tạo tâm cụm khác nhau
- ☐ c. Hiệu năng của thuật toán K-means có thể phụ thuộc vào chọn độ đo khoảng cách giữa 2 điểm
- ☐ d. Khi số lượng cụm tăng, lỗi huấn luyện sẽ có thể giảm
- ☐ e. K-means có thể chịu kết quả học tồi (underfitting) khi tăng số lượng cụm

ge

Next p



Time left 0:10:09

Question 6

Answer saved

Marked out of  
1.00

Giảm chiều dữ liệu có thể được sử dụng như một cách tiền xử lý cho các thuật toán học máy như cây quyết định, K-means, mạng nơ-ron, etc?

Select one:

☒ True

☐ False

Previous page

Next page

# n Học máy và khai phá dữ liệu

Question **13**

Not yet  
answered

Marked out of  
1.00

Một biến quan sát được có thể biểu diễn

Select one:

- ☐ a. bất kể thứ gì
- ☐ b. một sự kiện mà không thể quan sát được
- ☐ c. một sự kiện mà có thể quan sát được hoặc không
- ☒ d. một sự kiện mà có thể quan sát được trong thực tế

[Previous page](#)

# máy và khai phá dữ liệu

Những phát biểu nào sau đây là SAI khi huấn luyện bộ phân loại SVM tuyến tính từ một tập huấn luyện kích thước  $m$ ?

Select one or more:

- ☐ a. Số lượng vector hỗ trợ độc lập với  $m$
- ☒ b. Bộ phân loại học được có thể đưa ra dự đoán sai cho tất cả tập dữ liệu huấn luyện
- ☒ c. Số lượng vector hỗ trợ phụ thuộc vào  $m$
- ☐ d. Số lượng vector hỗ trợ phụ thuộc vào chọn hằng số phạt  $C$
- ☒ e. Có thể không có vector hỗ trợ



## Thần Học máy và khai phá dữ liệu

Question 3

Not yet  
answered

Marked out of  
1.00

Tất cả các mạng nơ-ron tính toán các hàm phi tuyến theo bộ tham số của nó?

Select one:

☒ True

☐ False

[Previous page](#)

Time left 0:06:01

Question 38

Answer saved

Marked out of  
1.00

Giải thuật học cây quyết định ID3 ...

Select one:

- ☒ a. có thể dùng cho bài toán hồi quy
- ✓ ☐ b. có thể dùng cho bài toán phân loại
- ☐ c. có thể dùng cho cả bài toán phân loại và bài toán hồi quy
- ☐ d. không thể dùng cho bài toán phân loại cũng như bài toán hồi quy

Clear my choice

Previous page

Next page

... và ...

Song cho biết hai quy biến ngẫu nhiên  $X$  và  $Y$  có hàm mật độ xác suất theo đầu vào  $x$ .

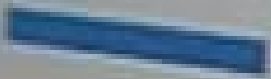
Select one:

☒ a. Đúng

☐ b. Sai

Check my answers

Time left 00:05:44





Question 35

Answer saved

Marked out of  
1.00

Rời rạc hoá dữ liệu (Data discretization) là

Select one:

- ☒ a. chuyển đổi từ liên tục thành rời rạc
- ☐ b. chuyển đổi dữ liệu từ rời rạc thành liên tục
- ☐ c. tìm loại sai (outliers)
- ☐ d. mở rộng dữ liệu (scale up data)

[Clear my choice](#)[Previous page](#)[Next page](#)

Question 35

Answer saved

Marked out of  
1.00

Rời rạc hoá dữ liệu (Data discretization) là

Select one:

- ☒ a. chuyển đổi từ liên tục thành rời rạc
- ☐ b. chuyển đổi dữ liệu từ rời rạc thành liên tục
- ☐ c. tìm loại sai (outliers)
- ☐ d. mở rộng dữ liệu (scale up data)

[Clear my choice](#)[Previous page](#)[Next page](#)

# háy và khai phá dữ liệu

Time left 03

Những hàm khoảng cách nào sau đây có thể được dùng trong giải thuật học láng giềng gần nhất đối với các biến (thuộc tính) kiểu định danh

- ✓ 1. Hàm khoảng cách Hamming
- ✓ 2. Hàm khoảng cách Manhattan
- 3. Hàm khoảng cách Euclidean

Select one:

- ☐ a. 1, 2 and 3
- ☐ b. 1
- ☐ c. 3
- ☐ d. 2
- ☐ e. 1 and 2
- ☐ f. 2 and 3
- ☐ g. 1 and 3

page

Next page



Time left 0:06:59

Question 33

Answer saved

Marked out of  
1.00

Thuật toán Perceptron sẽ hội tụ

Select one:

- ☐ a. luôn luôn hội tụ
- ☐ b. dù cho dữ liệu không phân tách tuyến tính
- ☒ c. nếu dữ liệu phân tách tuyến tính
- ☐ d. nếu chúng ta khởi tạo tất cả các tham số bằng 0

Clear my choice

Previous page

Next page

Time left 0:06:59

Question 33

Answer saved

Marked out of  
1.00

Thuật toán Perceptron sẽ hội tụ

Select one:

- ☐ a. luôn luôn hội tụ
- ☐ b. dù cho dữ liệu không phân tách tuyến tính
- ☒ c. nếu dữ liệu phân tách tuyến tính
- ☐ d. nếu chúng ta khởi tạo tất cả các tham số bằng 0

Clear my choice

Previous page

Next page

Time left 0:10:20

Question 4

Answer saved

Marked out of  
1.00

Tiền xử lý dữ liệu có thể xem như

Select one:

- ☐ a. là quá trình quyết định định dạng về dữ liệu và thu thập dữ liệu
- ☒ b. là quá trình biến đổi dữ liệu thô tới một dạng thích hợp để cho bước phân tích
- ☐ c. phân tích một số tích chất chính của tập dữ liệu

Clear my choice

Previous page

Next page

Time left 0:07:46

Question 27

Answer saved

Marked out of  
1.00

Cây quyết định ...

Select one:

- ☒ a. là một phương pháp học có giám sát
- ☐ b. không phải là phương pháp học có giám sát và cũng không phải là phương pháp học không giám sát
- ☐ c. là một phương pháp học không giám sát
- ☐ d. là một phương pháp học có giám sát và cũng là một phương pháp học không giám sát

[Clear my choice](#)

[Previous page](#)

[Next page](#)



Time left 0:07:46

Question 27

Answer saved

Marked out of  
1.00

Cây quyết định ...

Select one:

- ✓ ☐ a. là một phương pháp học có giám sát
- ☒ b. không phải là phương pháp học có giám sát và cũng không phải là phương pháp học không giám sát
- ☐ c. là một phương pháp học không giám sát
- ☐ d. là một phương pháp học có giám sát và cũng là một phương pháp học không giám sát

Clear my choice

Previous page

Next page

Time left 0:00:29

Question 19

Answer saved

Marked out of  
1.00

Trong khai phá luật kết hợp, tổng số hỗ trợ (support count) của tập mục (itemset) A được định nghĩa là gì?

Select one:

- ☒ a. Tỷ lệ tổng số giao dịch chứa A và tổng số tất cả giao dịch
- ☒ b. Tổng số lượng giao dịch chứa A
- ☐ c. Tỷ lệ tổng số giao dịch không chứa A và tổng số tất cả giao dịch
- ☐ d. Tổng số giao dịch không chứa A

[Clear my choice](#)

[Previous page](#)

[Next page](#)

Time left 0:08:18

Question 21

Answer saved

Marked out of  
1.00

Tất cả các mạng nơ-ron tính toán các hàm phi tuyến theo bộ tham số của nó?

Select one:

☒ True

☐ False

Previous page

Next page

Time left 0:08:08

## Question 22

Answer saved

Marked out of  
1.00

Giả sử có ma trận nhầm lẫn dưới đây. Độ chính xác của bộ phân loại là bao nhiêu nếu nó phân đoán một người có mua máy tính hay không?

	<i>Buy_computer = yes</i>	<i>Buy_computer = no</i>	<i>Total</i>
<i>Buy_computer = yes</i>	6954	46	7000
<i>Buy_computer = no</i>	412	2588	3000
<i>Total</i>	7366	2634	10000

Select one:

- ☐ a.  $(6954 + 46)/10000$
- ☐ b.  $(6954 + 412)/10000$
- ☒ c.  $(6954 + 2588)/10000$
- ☐ d.  $(46 + 412)/10000$



Time left 0:08:08

## Question 22

Answer saved

Marked out of  
1.00

Giả sử có ma trận nhầm lẫn dưới đây. Độ chính xác của bộ phân loại là bao nhiêu nếu nó phân đoán một người có mua máy tính hay không?

	<i>Buy_computer = yes</i>	<i>Buy_computer = no</i>	<i>Total</i>
<i>Buy_computer = yes</i>	6954	46	7000
<i>Buy_computer = no</i>	412	2588	3000
<i>Total</i>	7366	2634	10000

Select one:

- ☐ a.  $(6954 + 46)/10000$
- ☐ b.  $(6954 + 412)/10000$
- ☒ c.  $(6954 + 2588)/10000$
- ☐ d.  $(46 + 412)/10000$

# ọc máy và khai phá dữ liệu

Question **18**

Not yet  
answered

Marked out of  
1.00

Thuật toán K-means cho phân cụm đảm bảo hội tụ tới một điểm cực trị địa phương hoặc điểm dừng?

Select one:

- ☒ True  
☐ False

[Previous page](#)

Question 39

Answer saved

Marked out of  
1.00

Time left 0:05:51

Giả sử rằng chúng ta có một cơ sở dữ liệu trong đó mỗi giao dịch chứa cả tên và giá của từng item. Thông tin về giá có ảnh hưởng đến chất lượng của thuật toán Apriori không?

Select one:

☒ True

✓ ☐ False

Previous page

Next page

Question 39

Answer saved

Marked out of  
1.00

Time left 0:05:51

Giả sử rằng chúng ta có một cơ sở dữ liệu trong đó mỗi giao dịch chứa cả tên và giá của từng item. Thông tin về giá có ảnh hưởng đến chất lượng của thuật toán Apriori không?

Select one:

☒ True

☒ False

Previous page

Next page



# máy và khai phá dữ liệu

Tìm

5 Nhắc lại hàm mục tiêu của soft-margin SVM có dạng  $\|w\|^2 + C \sum_i \xi_i$ . Giảm giá trị của C có thể dẫn đến

Select one:

- ☐ a. giảm khả năng overfitting
- ☒ b. ít nhạy cảm với ngoại lai (outliers)
- ☐ c. lề nhỏ hơn
- ☐ d. tăng khả năng overfitting

Time left 0:09:54

Question 9

Answer saved

Marked out of  
1.00

Tại sao bước tìm tập mục thường xuyên với kích thước 2 (2-itemsets) thường là khó nhất?

Select one:

- ☐ a. Đây luôn là bước bắt buộc để tìm bất kỳ tập mục thường xuyên lớn hơn.
- ☒ b. Với một xác suất cao, nó yêu cầu bộ nhớ lớn nhất so với các loại tập thường xuyên kích thước khác.
- ☒ c. Lý do chính là thời gian tính toán lớn

Clear my choice

Previous page

Next page

## Machine Learning và khai phá dữ liệu

Question **44**

Not yet  
answered

Marked out of  
1.00

Với soft-margin SVM, những mẫu nào sẽ có giá trị  $\xi_i$  khác 0

Select one or more:

- ☒ a. Tất cả các mẫu nằm bên trong lề
- ☐ b. Tất cả các mẫu nằm trên siêu phẳng lề
- ☒ c. Tất cả các mẫu bị sai lớp
- ☐ d. Tất cả các mẫu nằm ngoài lề

[Previous page](#)