

Hướng dẫn:

- Sinh viên làm Đề tương ứng với **ký số cuối cùng trong MSSV** của mình.
- Thời gian làm bài là 60 phút
- Được sử dụng tất cả mọi tài liệu, kể cả Internet

Ví dụ:

Sinh viên Nguyễn Văn A có MSSV là 018123456 sẽ làm đề số 6.

Đề 1

Phần 1. Câu hỏi ngắn

Trả lời/chọn câu trả lời đúng trong các câu quiz bên dưới và giải thích ngắn gọn cho câu trả lời của mình.

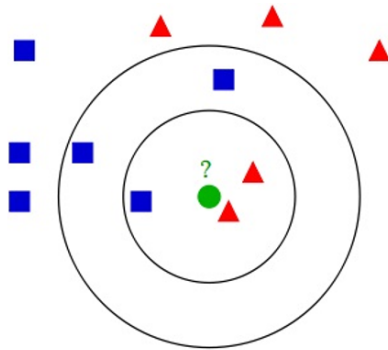
Câu 1 : Đặc tính nào là đúng đối với Supervised Learning?

- a. Labeled data.
- b. Direct feedback.
- c. Predict outcome/future
- d. All are correct

Câu 2 : Thứ tự nào là đúng đối với mô hình Classification?

- a. Classification Model => Training Data => Make Prediction on unseen Data
- b. Training Data => Classification Model => Make Prediction on unseen Data
- c. Make Prediction on unseen Data => Training Data => Classification Model
- d. Training Data => Make Prediction on unseen Data => Classification Model

Câu 3 : Chọn câu trả lời sai:



- a. Khi $k = 1$, điểm màu xanh thuộc về lớp màu đỏ
- b. Khi $k = 3$, điểm màu xanh thuộc về lớp màu đỏ
- c. Khi $k = 5$, điểm màu xanh thuộc về lớp màu xanh
- d. Tất cả các câu khác đều sai

Phần 2. Tự luận

Câu 1.

Kỹ thuật học máy đã sử dụng trong bài tập lớn của bạn là gì. Ưu và khuyết điểm của giải thuật này.

Câu 2.

Vẽ state-chart diagram của hệ thống đã phát triển trong bài tập lớn của bạn.

Đề 2

Phần 1. Câu hỏi ngắn

Trả lời/chọn câu trả lời đúng trong các câu quiz bên dưới và giải thích ngắn gọn cho câu trả lời của mình.

Câu 1 : Cho ma trận confusion matrix như bên dưới. Biết rằng 0 là positive và 1 là negative, giá trị Precision bằng bao nhiêu?

Classification Confusion Matrix		
	Predicted Class	
Actual Class	1	0
1	201	85
0	25	2689

- a. 0.89
- b. 0.96
- c. 0.90
- d. 0.80

Câu 2 : Cho ma trận confusion matrix như bên dưới. Biết rằng 0 là positive và 1 là negative, giá trị Recall bằng bao nhiêu?

Classification Confusion Matrix		
	Predicted Class	
Actual Class	1	0
1	201	85
0	25	2689

- a. 0.99
- b. 0.86
- c. 0.80
- d. None of above

Câu 3 : Câu phát biểu nào là đúng đối với Decision Tree?

- a. Mỗi nút trung gian để kiểm tra một thuộc tính
- b. Mỗi cạnh ứng với một giá trị thuộc tính
- c. Mỗi nút lá ứng với một giá trị phân loại (nhãn)
- d. Tất cả đều đúng

Phần 2. Tự luận

Câu 1.

Kỹ thuật học máy đã sử dụng trong bài tập lớn của bạn là gì. Ưu và khuyết điểm của giải thuật này.

Câu 2.

Vẽ state-chart diagram của hệ thống đã phát triển trong bài tập lớn của bạn.

Đề 3

Phần 1. Câu hỏi ngắn

Trả lời/chọn câu trả lời đúng trong các câu quiz bên dưới và giải thích ngắn gọn cho câu trả lời của mình.

Câu 1 : Giả sử chúng ta có một tập dữ liệu bao gồm 1000 mẫu, trong đó số mẫu $\text{income}=\text{low}$ (0), $\text{income}=\text{medium}$ (990), and $\text{income}=\text{high}$ (10). Nếu chúng ta dùng kỹ thuật Laplacian correction và tính xác suất $\text{Prob}(\text{income}=\text{low})$ sau đó thì giá trị tương ứng sẽ là

- a. $\text{Prob}(\text{income}=\text{low}) = 1/1003$
- b. $\text{Prob}(\text{income}=\text{low}) = 1/1000$
- c. $\text{Prob}(\text{income}=\text{low}) = 1/900$
- d. All are not corrects.

Câu 2 :

Câu phát biểu nào không đúng với cluster analysis?

Cluster analysis tương ứng với...

- a. Supervised classification
- b. Simple segmentation
- c. Result of a query
- d. All are correct.

Câu 3 : Phát biểu nào về giải thuật k -means là đúng

- a. Đây là giải thuật theo hướng partitional clustering
- b. Mỗi cluster ứng với một centroid (center point)
- c. Ở mỗi bước lặp, mỗi điểm sẽ được gán với cluster có centroid gần nhất
- d. Số cluster, tức k , cần phải được xác định trước
- e. Tất cả đều đúng

Phần 2. Tự luận

Câu 1.

Kỹ thuật học máy đã sử dụng trong bài tập lớn của bạn là gì. Ưu và khuyết điểm của giải thuật này.

Câu 2.

Vẽ state-chart diagram của hệ thống đã phát triển trong bài tập lớn của bạn.

Đề 4

Phần 1. Câu hỏi ngắn

Trả lời/chọn câu trả lời đúng trong các câu quiz bên dưới và giải thích ngắn gọn cho câu trả lời của mình.

Câu 1 : Đặc tính nào là đúng đối với Supervised Learning?

- a. Labeled data.
- b. Direct feedback.
- c. Predict outcome/future
- d. All are correct

Câu 2 : Cho ma trận confusion matrix như bên dưới. Biết rằng 0 là positive và 1 là negative, giá trị Precision bằng bao nhiêu?

Classification Confusion Matrix		
	Predicted Class	
Actual Class	1	0
1	201	85
0	25	2689

- a. 0.89
- b. 0.96
- c. 0.90
- d. 0.80

Câu 3 : Giả sử chúng ta có một tập dữ liệu bao gồm 1000 mẫu, trong đó số mẫu income=low (0), income= medium (990), and income = high (10). Nếu chúng ta dùng kỹ thuật Laplacian correction và tính xác suất $\text{Prob}(\text{income} = \text{low})$ sau đó thì giá trị tương ứng sẽ là

- a. $\text{Prob}(\text{income} = \text{low}) = 1/1003$
- b. $\text{Prob}(\text{income} = \text{low}) = 1/1000$
- c. $\text{Prob}(\text{income} = \text{low}) = 1/900$
- d. All are not corrects.

Phần 2. Tự luận

Câu 1.

Kỹ thuật học máy đã sử dụng trong bài tập lớn của bạn là gì. Ưu và khuyết điểm của giải thuật này.

Câu 2.

Vẽ state-chart diagram của hệ thống đã phát triển trong bài tập lớn của bạn.

Đề 5

Phần 1. Câu hỏi ngắn

Trả lời/chọn câu trả lời đúng trong các câu quiz bên dưới và giải thích ngắn gọn cho câu trả lời của mình.

Câu 1 : Thứ tự nào là đúng đối với mô hình Classification?

- a. Classification Model => Training Data => Make Prediction on unseen Data
- b. Training Data => Classification Model => Make Prediction on unseen Data
- c. Make Prediction on unseen Data => Training Data => Classification Model
- d. Training Data => Make Prediction on unseen Data => Classification Model

Câu 2 : Cho ma trận confusion matrix như bên dưới. Biết rằng 0 là positive và 1 là negative, giá trị Recall bằng bao nhiêu?

Classification Confusion Matrix		
	Predicted Class	
Actual Class	1	0
1	201	85
0	25	2689

- a. 0.99
- b. 0.86
- c. 0.80
- d. None of above

Câu 3 :

Câu phát biểu nào không đúng với cluster analysis?

Cluster analysis tương ứng với...

- a. Supervised classification
- b. Simple segmentation
- c. Result of a query
- d. All are correct.

Phần 2. Tự luận

Câu 1.

Kỹ thuật học máy đã sử dụng trong bài tập lớn của bạn là gì. Ưu và khuyết điểm của giải thuật này.

Câu 2.

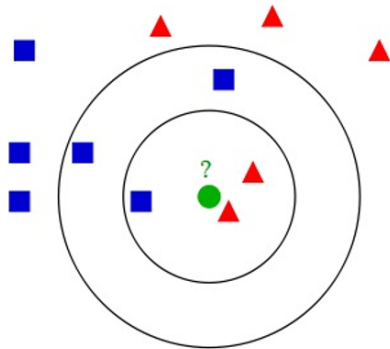
Vẽ state-chart diagram của hệ thống đã phát triển trong bài tập lớn của bạn

Đề 6

Phần 1. Câu hỏi ngắn

Trả lời/chọn câu trả lời đúng trong các câu quiz bên dưới và giải thích ngắn gọn cho câu trả lời của mình.

Câu 1 : Chọn câu trả lời sai:



- a. Khi $k = 1$, điểm màu xanh thuộc về lớp màu đỏ
- b. Khi $k = 3$, điểm màu xanh thuộc về lớp màu đỏ
- c. Khi $k = 5$, điểm màu xanh thuộc về lớp màu xanh
- d. Tất cả các câu khác đều sai

Câu 2 : Câu phát biểu nào là đúng đối với Decision Tree?

- a. Mỗi nút trung gian để kiểm tra một thuộc tính
- b. Mỗi cạnh ứng với một giá trị thuộc tính
- c. Mỗi nút lá ứng với một giá trị phân loại (nhãn)
- d. Tất cả đều đúng

Câu 3 : Phát biểu nào về giải thuật k -means là đúng

- a. Đây là giải thuật theo hướng partitional clustering
- b. Mỗi cluster ứng với một centroid (center point)
- c. Ở mỗi bước lặp, mỗi điểm sẽ được gán với cluster có centroid gần nhất
- d. Số cluster, tức k , cần phải được xác định trước
- e. Tất cả đều đúng

Phần 2. Tự luận

Câu 1.

Kỹ thuật học máy đã sử dụng trong bài tập lớn của bạn là gì. Ưu và khuyết điểm của giải thuật này.

Câu 2.

Vẽ state-chart diagram của hệ thống đã phát triển trong bài tập lớn của bạn

Đề 7

Phần 1. Câu hỏi ngắn

Trả lời/chọn câu trả lời đúng trong các câu quiz bên dưới và giải thích ngắn gọn cho câu trả lời của mình.

Câu 1 : Đặc tính nào là đúng đối với Supervised Learning?

- a. Labeled data.
- b. Direct feedback.
- c. Predict outcome/future
- d. All are correct

Câu 2 : Câu phát biểu nào là đúng đối với Decision Tree?

- a. Mỗi nút trung gian để kiểm tra một thuộc tính
- b. Mỗi cạnh ứng với một giá trị thuộc tính
- c. Mỗi nút lá ứng với một giá trị phân loại (nhãn)
- d. Tất cả đều đúng

Câu 2 :

Câu phát biểu nào không đúng với cluster analysis?

Cluster analysis tương ứng với...

- a. Supervised classification
- b. Simple segmentation
- c. Result of a query
- d. All are correct.

Phần 2. Tự luận

Câu 1.

Kỹ thuật học máy đã sử dụng trong bài tập lớn của bạn là gì. Ưu và khuyết điểm của giải thuật này.

Câu 2.

Vẽ state-chart diagram của hệ thống đã phát triển trong bài tập lớn của bạn

Đề 8

Phần 1. Câu hỏi ngắn

Trả lời/chọn câu trả lời đúng trong các câu quiz bên dưới và giải thích ngắn gọn cho câu trả lời của mình.

Câu 1 : Thứ tự nào là đúng đối với mô hình Classification?

- a. Classification Model => Training Data => Make Prediction on unseen Data
- b. Training Data => Classification Model => Make Prediction on unseen Data
- c. Make Prediction on unseen Data => Training Data => Classification Model
- d. Training Data => Make Prediction on unseen Data => Classification Model

Câu 2 : Phát biểu nào về giải thuật k -means là đúng

- a. Đây là giải thuật theo hướng partitional clustering
- b. Mỗi cluster ứng với một centroid (center point)
- c. Ở mỗi bước lặp, mỗi điểm sẽ được gán với cluster có centroid gần nhất
- d. Số cluster, tức k , cần phải được xác định trước
- e. Tất cả đều đúng

Câu 3 : Cho ma trận confusion matrix như bên dưới. Biết rằng 0 là positive và 1 là negative, giá trị Precision bằng bao nhiêu?

Classification Confusion Matrix		
	Predicted Class	
Actual Class	1	0
1	201	85
0	25	2689

- a. 0.89
- b. 0.96
- c. 0.90
- d. 0.80

Phần 2. Tự luận

Câu 1.

Kỹ thuật học máy đã sử dụng trong bài tập lớn của bạn là gì. Ưu và khuyết điểm của giải thuật này.

Câu 2.

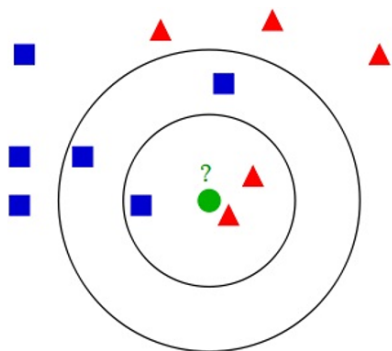
Vẽ state-chart diagram của hệ thống đã phát triển trong bài tập lớn của bạn

Đề 9

Phần 1. Câu hỏi ngắn

Trả lời/chọn câu trả lời đúng trong các câu quiz bên dưới và giải thích ngắn gọn cho câu trả lời của mình.

Câu 1 : Chọn câu trả lời sai:



- a. Khi $k = 1$, điểm màu xanh thuộc về lớp màu đỏ
- b. Khi $k = 3$, điểm màu xanh thuộc về lớp màu đỏ
- c. Khi $k = 5$, điểm màu xanh thuộc về lớp màu xanh
- d. Tất cả các câu khác đều sai

Câu 2 : Cho ma trận confusion matrix như bên dưới. Biết rằng 0 là positive và 1 là negative, giá trị Precision bằng bao nhiêu?

Classification Confusion Matrix		
Actual Class	Predicted Class	
	1	0
1	201	85
0	25	2689

- a. 0.89
- b. 0.96
- c. 0.90
- d. 0.80

Câu 3 :

Câu phát biểu nào không đúng với cluster analysis?

Cluster analysis tương ứng với...

- a. Supervised classification
- b. Simple segmentation
- c. Result of a query
- d. All are correct.

Phần 2. Tự luận

Câu 1.

Kỹ thuật học máy đã sử dụng trong bài tập lớn của bạn là gì. Ưu và khuyết điểm của giải thuật này.

Câu 2.

Vẽ state-chart diagram của hệ thống đã phát triển trong bài tập lớn của bạn

Đề 10

Phần 1. Câu hỏi ngắn

Trả lời/chọn câu trả lời đúng trong các câu quiz bên dưới và giải thích ngắn gọn cho câu trả lời của mình.

Câu 1 : Thứ tự nào là đúng đối với mô hình Classification?

- a. Classification Model => Training Data => Make Prediction on unseen Data
- b. Training Data => Classification Model => Make Prediction on unseen Data
- c. Make Prediction on unseen Data => Training Data => Classification Model
- d. Training Data => Make Prediction on unseen Data => Classification Model

Câu 2 : Câu phát biểu nào là đúng đối với Decision Tree?

- a. Mỗi nút trung gian để kiểm tra một thuộc tính
- b. Mỗi cạnh ứng với một giá trị thuộc tính
- c. Mỗi nút lá ứng với một giá trị phân loại (nhãn)
- d. Tất cả đều đúng

Câu 3 : Giả sử chúng ta có một tập dữ liệu bao gồm 1000 mẫu, trong đó số mẫu income=low (0), income= medium (990), and income = high (10). Nếu chúng ta dùng kỹ thuật Laplacian correction và tính xác suất $\text{Prob}(\text{income} = \text{low})$ sau đó thì giá trị tương ứng sẽ là

- a. $\text{Prob}(\text{income} = \text{low}) = 1/1003$
- b. $\text{Prob}(\text{income} = \text{low}) = 1/1000$
- c. $\text{Prob}(\text{income} = \text{low}) = 1/900$
- d. All are not corrects.

Phần 2. Tự luận

Câu 1.

Kỹ thuật học máy đã sử dụng trong bài tập lớn của bạn là gì. Ưu và khuyết điểm của giải thuật này.

Câu 2.

Vẽ state-chart diagram của hệ thống đã phát triển trong bài tập lớn của bạn