

# Final Exam

**Due Dec 9 at 1:15am      Points 40      Questions 20**

**Available** Dec 6 at 12am - Dec 9 at 2am 3 days

**Time Limit** 75 Minutes

## Instructions

No collaboration

75 mins

## Attempt History

	<b>Attempt</b>	<b>Time</b>	<b>Score</b>
<b>LATEST</b>	<a href="#"><u>Attempt 1</u></a>	41 minutes	40 out of 40

! Correct answers are hidden.

Score for this quiz: **40** out of 40

Submitted Dec 8 at 6:17pm

This attempt took 41 minutes.

<b>Question 1</b>	<b>2 / 2 pts</b>
Apriori pruning based on support is a greedy strategy but is not optimal	
<input type="radio"/> True	
<input checked="" type="radio"/> False	

**Question 2****2 / 2 pts**

Which of the above statements are true for any A, B, and C?

- If  $A \rightarrow B$  then  $B \rightarrow A$ .
- If  $A \rightarrow B$  and  $B \rightarrow C$  then  $A \rightarrow C$ .
- If  $A \rightarrow C$  then  $A \cup B \rightarrow C$ .
- If  $A \cup B \rightarrow C$  then  $A \rightarrow C$ .
- I & II
- I, II, & III
- I, II, & IV
- II & III
- none

**Question 3****2 / 2 pts**

Example of market basket transactions.

Transaction ID Items Bought

- |    |              |
|----|--------------|
| 1  | {a, b, d, e} |
| 2  | {b, c, d}    |
| 3  | {a, b, d, e} |
| 4  | {a, c, d, e} |
| 5  | {b, c, d, e} |
| 6  | {b, d, e}    |
| 7  | {c, d}       |
| 8  | {a, b, c}    |
| 9  | {a, d, e}    |
| 10 | {b, d}       |

Compute the confidence for the association rules {a, e} -> {b}.

0.5

#### Question 4

2 / 2 pts

Example of market basket transactions.

**Transaction ID Items Bought**

1 {a, b, d, e}

2 {b, c, d}

3 {a, b, d, e}

4 {a, c, d, e}

5 {b, c, d, e}

6 {b, d, e}

7 {c, d}

8 {a, b, c}

9 {a, d, e}

10 {b, d}

If the minimum support is set at 40%, how many frequent 3-itemsets will be found? Note that if a 3 itemset is a subset of a larger itemset, it counts as one occurrence.

2

**Question 5****2 / 2 pts**

## Example of market basket transactions.

### Transaction ID Items Bought

1 {a, b, d, e}

2 {b, c, d}

3 {a, b, d, e}

4 {a, c, d, e}

5 {b, c, d, e}

6 {b, d, e}

7 {c, d}

8 {a, b, c}

9 {a, d, e}

10 {b, d}

Compute the support for itemset {b,c,d}.

0.2

**Question 6****2 / 2 pts**

Which of the following is an application of autoencoders?

- 
- feature learning.
- 
- dimensionality reduction.
- 
- Rule mining
- 
- all of the above.

**Question 7****2 / 2 pts**

What is an autoencoder?

- 
- a neural network that copies its input to its output.
- 
- a neural network that codes itself.
- 
- 
- a neural network that maps an output to an input through a hidden layer.
- 
- 
- a neural network that is trained to attempt to copy its input to its output.

**Question 8****2 / 2 pts**

The number of neurons in the output layer should match the number of classes (Where the number of classes is greater than 2) in a supervised learning task.

 True False**Question 9****2 / 2 pts**

Select all that is true about Restricted Boltzman Machine (RBM)

 Can be used for recommendation systems RBM training is often probabilistic RBM is a kind of Deep Belief Network None of the above.**Question 10****2 / 2 pts**

Consider a movie recommendation application, where an user can log in and select movies they like based on their previous preferences.

Which neural network architecture would be suitable to complete this task

- Fully-Connected Neural Network.
- Convolutional Neural Network.
- Recurrent Neural Network.
- Restricted Boltzmann Machine.

### Question 11

2 / 2 pts

Review the following table.

Data	X	Y
x1	1	1
x2	6	6
x3	7	7
x4	3	2
x5	6	8
x6	2	1
x7	2	3

Compute the average silhouette score to evaluate K-mean clustering result with the parameters below.

K:2

Initial centroids: centroid1 = (4,3) and centroid2 = (9,7)

Distance measure: Euclidean distance

0.752

### Question 12

2 / 2 pts

Review the following table.

Data	X	Y
A	1	2
B	2	2
C	2	3
D	8	7
E	8	8
F	8	2

Select the correct number of core points and DBSCAN clustering result with the parameters below.

Eps: 5

Minimum samples (MinPts):3

Distance measurement: Euclidean distance

- 5, cluster1: {A,B,C}, cluster2: {D,E,F}
- 6, cluster1: {A,B}, cluster2: {C,D,E}, cluster3: {F}
- 4, cluster1: {A,B,C}, cluster2: {D, E, F}



5, cluster1: {A}, cluster2: {B}, cluster3: {C}, cluster4: {D}, cluster5: {E}



None of above

### Question 13

2 / 2 pts

Review the following table.

Data	X	Y
x1	1	1
x2	6	6
x3	7	7
x4	3	2
x5	6	8
x6	2	1
x7	2	3

Compute the sum of squared error (SSE) to evaluate K-mean clustering result with the parameters below.

K:2

Initial centroids: centroid1 = (4,3) and centroid2 = (9,7)

Distance measure: Euclidean distance

7.4167

**Question 14****2 / 2 pts**

Select different aspects of cluster validation.

- Determining the clustering tendency of a set of data
- To improve the complexity of the algorithm
- Determining the 'correct' number of clusters
- None of the above

correct these are some of the aspects of cluster validation!

**Question 15****2 / 2 pts**

Select qualities of clusters produced by a good clustering algorithm.

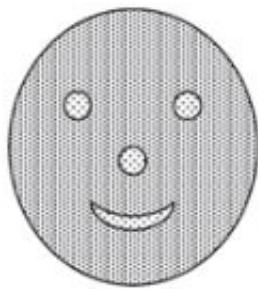
- Intra-cluster distances are minimized
- Inter-cluster distances are maximized
- Number of clusters produced
- A and B are correct

good cluster have strong cohesion within the cluster and maximum distance between them.

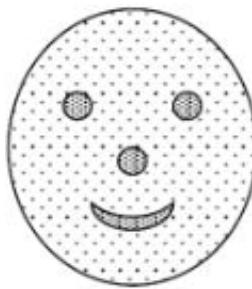
### Question 16

2 / 2 pts

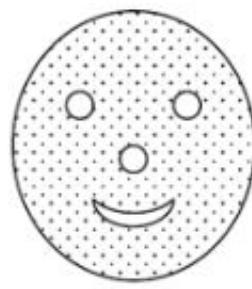
Review the image below.



(a)



(b)



(c)



(d)

**Darkness or number of dots represent density. Lines are used only to distinguish regions and do not represent points.**

If we want to find the patterns represented by the nose, eyes, and mouth using single linkage hierarchical clustering, select all figures that be well-clustered?

(a)

(b)

(c)

(d)

- None of above

**Question 17****2 / 2 pts**

Select all that are not true about DBSCAN.

- it has trouble when the clusters have widely varying densities
- it is very sensitive to noise and to the size of the data
- DBSCAN has a complexity of order  $O(n^3)$
- 

it can be expensive when the computation of nearest neighbors requires computing all pairwise proximities

- A and B are correct
- B and C are correct

DBSCAN is relatively resistant to noise and can handle clusters of arbitrary shapes and sizes. It has complexity order  $O(n \log n)$

- A, C and D are correct

**Question 18****2 / 2 pts**

### Select the true statement for Nearest Neighbor classification



Although Nearest Neighbor classification has the data dimensionality issue, it can be solved by the scaling technique



k-NN classifier is a typical lazy learner because it spend so much time building the model even if the 'k' is very small



Drawing a circle is a mandatory task to find neighbor classes regardless of the number of 'K'



Determining the optimal value of 'K' is important for k-NN classifier performance.



None of above

### Question 19

2 / 2 pts

Review the table below.

Rainy	Sunshine		Wind		Atmosphere		Temperature	
	Yes	No	Yes	No	High	Low	High	Low
Yes	2	6	6	2	8	0	5	3
No	8	4	2	10	2	10	6	6

Is it raining if the weather condition is *sunshine, no wind, high atmosphere, and low temperature* (Hint: you should use Naïve Bayes)?

Yes

No

## Question 20

2 / 2 pts

Review the table below.

A	B	C	Class
F	F	F	Y
F	F	T	N
F	T	T	N
F	T	T	N
F	F	T	Y
T	F	T	Y
T	F	T	N
T	F	T	N
T	T	T	Y
T	F	T	Y

Select the correct prediction class for a test sample (A=F, B=T, C=F) using the naïve Bayes approach.

---

'Y'

---

'N'

---

Anyone

---

Quiz Score: **40** out of 40

# Final Exam

**Due** Dec 9 at 1:15am      **Points** 40      **Questions** 20

**Available** Dec 6 at 12am - Dec 9 at 2am 3 days

**Time Limit** 75 Minutes

## Instructions

No collaboration

75 mins

## Attempt History

	<b>Attempt</b>	<b>Time</b>	<b>Score</b>
LATEST	<a href="#">Attempt 1</a>	43 minutes	38 out of 40

ⓘ Correct answers are hidden.

Score for this quiz: **38** out of 40

Submitted Dec 8 at 5:27pm

This attempt took 43 minutes.

### Question 1

2 / 2 pts

Transaction ID Items Bought

- 1 {Milk, Beer, Diapers}
- 2 {Bread, Butter, Milk}
- 3 {Milk, Diapers, Cookies}
- 4 {Bread, Butter, Cookies}

- 5      {Beer, Cookies, Diapers}
- 6      {Milk, Diapers, Bread, Butter}
- 7      {Bread, Butter, Diapers}
- 8      {Beer, Diapers}
- 9      {Milk, Diapers, Bread, Butter}
- 10     {Beer, Cookies}

What is the confidence of the rule Beer  $\rightarrow$  Cookies

0.5

## Question 2

2 / 2 pts

Transaction ID Items Bought

- 1      {Milk, Beer, Diapers}
- 2      {Bread, Butter, Milk}
- 3      {Milk, Diapers, Cookies}
- 4      {Bread, Butter, Cookies}
- 5      {Beer, Cookies, Diapers}
- 6      {Milk, Diapers, Bread, Butter}

- 7      {Bread, Butter, Diapers}
- 8      {Beer, Diapers}
- 9      {Milk, Diapers, Bread, Butter}
- 10     {Beer, Cookies}

What is the maximum number of size 3-itemsets that can be derived from this data set. Note that the itemsets derived may not be in the dataset.

- 
- 20
- 
- 35
- 
- 40
- 
- 15
- 
- None of the above.

### Question 3

2 / 2 pts

Example of market basket transactions.

Transaction ID Items Bought

- 1      {a, b, d, e}
- 2      {b, c, d}
- 3      {a, b, d, e}

- 4      {a, c, d, e}
- 5      {b, c, d, e}
- 6      {b, d, e}
- 7      {c, d}
- 8      {a, b, c}
- 9      {a, d, e}
- 10     {b, d}

If the minimum support is set at 40%, how many frequent 3-itemsets will be found? Note that if a 3 itemset is a subset of a larger itemset, it counts as one occurrence.

2

**Question 4****2 / 2 pts**

Transaction ID Items Bought

- 1      {Milk, Beer, Diapers}
- 2      {Bread, Butter, Milk}
- 3      {Milk, Diapers, Cookies}
- 4      {Bread, Butter, Cookies}

- 5      {Beer, Cookies, Diapers}
- 6      {Milk, Diapers, Bread, Butter}
- 7      {Bread, Butter, Diapers}
- 8      {Beer, Diapers}
- 9      {Milk, Diapers, Bread, Butter}
- 10     {Beer, Cookies}

Find all itemsets (of size 3 or larger) that has the largest support.

- 
- {Milk, Beer, Diapers}
- 
- {Bread, Butter, Milk}
- 
- {Bread, Butter, Diapers}
- 
- {Milk, Diapers, Bread}
- 
- None of the above.
- 
- All of the above.

### Question 5

2 / 2 pts

Example of market basket transactions.

Transaction ID Items Bought

- 1      {a, b, d, e}

- 2      {b, c, d}
- 3      {a, b, d, e}
- 4      {a, c, d, e}
- 5      {b, c, d, e}
- 6      {b, d, e}
- 7      {c, d}
- 8      {a, b, c}
- 9      {a, d, e}
- 10     {b, d}

Compute the confidence for the association rules  $\{a, e\} \rightarrow \{b\}$ .

0.5

## Question 6

2 / 2 pts

Usage of deep learning is predicated with availability of large amounts of data. Is the data volume the only determining factor to check whether we can use deep learning ?

Yes, lower volumes lead to overfitting



No, balance among classes in the input data is the only determining factor  
volume does not matter



No, balance among classes in the input data is a co-factor along with the data volume



None of the above

### Question 7

2 / 2 pts

Select all that is true about Restricted Boltzman Machine (RBM)

Can be used for recommendation systems

RBM training is often probabilistic

RBM is a kind of Deep Belief Network

None of the above.

### Question 8

2 / 2 pts

Which of the following is an application of autoencoders?

feature learning.

dimensionality reduction.

Rule mining

all of the above.

**Question 9****2 / 2 pts**

Increase in size of a convolutional hidden layer would not necessarily increase the performance of a convolutional network.

---

True

---

False

**Question 10****2 / 2 pts**

Select the true statements about neural network (NN) complexity.

---

CNN is always more complex than DNN

---

Adding more hidden layers increases complexity of a NN

---

An Autoencoder is always more complex than a CNN

---



Complexity metric defined for a NN may not make sense for other classification engines

**Question 11****2 / 2 pts**

Review the following table.

Data	X	Y
A	1	2
B	2	2

C	2	3
D	8	7
E	8	8
F	8	2

Select the correct number of core points and DBSCAN clustering result with the parameters below.

Eps: 5

Minimum samples (MinPts):3

Distance measurement: Euclidean distance

- 5, cluster1: {A,B,C}, cluster2: {D,E,F}
- 6, cluster1: {A,B}, cluster2: {C,D,E}, cluster3: {F}
- 4, cluster1: {A,B,C}, cluster2: {D, E, F}
- 5, cluster1: {A}, cluster2: {B}, cluster3: {C}, cluster4: {D}, cluster5: {E}
- None of above

### Question 12

2 / 2 pts

Select qualities of clusters produced by a good clustering algorithm.

- Intra-cluster distances are minimized
- Inter-cluster distances are maximized
- Number of clusters produced
- A and B are correct

good cluster have strong cohesion within the cluster and maximum distance between them.

**Question 13****2 / 2 pts**

What is the purpose of cluster analysis?

- To avoid finding patterns in noise
- To compare clustering algorithms
- To compare two sets of clusters
- all of the above
- None of the above

we do cluster validity to avoid clustering noise and find a suitable algorithim for our data

**Question 14****2 / 2 pts**

Review the following table.

Data	X	Y
x1	1	1
x2	6	6
x3	7	7
x4	3	2
x5	6	8
x6	2	1
x7	2	3

Select the correct pair of final centroid and clustering set using K-mean clustering with the parameters below.

K:2

Initial centroids: centroid1 = (4,3) and centroid2 = (9,7)

Distance measure: Euclidean distance



Cluster 1: (x1, x4, x6, x7) with centroid=(2, 1.75), Cluster 2: (x2, x3, x5) with centroid=(6.33, 7)



Cluster 1: (x1, x4, x6, x7) with centroid=(1.75, 2), Cluster 2: (x2, x3, x5) with centroid=(6.33, 7)



Cluster 1: (x1, x4, x6, x7) with centroid=(2, 1.75), Cluster 2: (x2, x3, x5) with centroid=(7, 6.33)



Cluster 1: (x1, x2, x4, x6, x7) with centroid=(2, 1.75), Cluster 2: (x3, x5) with centroid=(6.33, 7)



Cluster 1:  $(x_1, x_2, x_4, x_6, x_7)$  with centroid=(2, 1.75), Cluster 2:  $(x_3, x_5)$  with centroid=(7, 6.33)

**Question 15****2 / 2 pts**

Review the following table.

Data	X	Y
x1	1	1
x2	6	6
x3	7	7
x4	3	2
x5	6	8
x6	2	1
x7	2	3

Compute the sum of squared error (SSE) to evaluate K-mean clustering result with the parameters below.

K:2

Initial centroids: centroid1 = (4,3) and centroid2 = (9,7)

Distance measure: Euclidean distance

7.4167

**Question 16****2 / 2 pts**

Select all statements that are true for Hierarchical clustering.

Although it requires high storage, high computing is not required

Robust to noise

Robust to outliers

It is not easy to decide on the number of cluster because the complicated dendrogram disturbs this decision

None of above

Incorrect

### Question 17

0 / 2 pts

Review the following table.

Data	X	Y
x1	1	1
x2	6	6
x3	7	7
x4	3	2
x5	6	8
x6	2	1
x7	2	3

Compute the updated centroids after the first iteration using K-mean clustering with the parameters below.

K:2

Initial centroids: centroid1 = (4,3) and centroid2 = (9,7)

Distance measure: Euclidean

Type your answer in the same format as "centroid1 = (x1,y1) and centroid2 = (x2,y2)"

Each co-ordinate should be in the format A.BC, where A, B and C are integers. If you have 4 as one of the co-ordinates you should write 4.00. Please be mindful of the spaces and other formats.

centroid1 = (2.00,1.75) and centroid2 = (6.33,7.00)

### Question 18

2 / 2 pts

Review the table below.

Instance	X-Axis	Class
i	-3	-
ii	-2	-
iii	-1	+
iv	0	+
v	1	+
vi	2	-

Select all the support vector instances using Non-linear SVM

Feature Map ( $\{R^1 \rightarrow R^2\}$ )

i

ii iii iv v vi**Question 19****2 / 2 pts**

Review the table below.

A	B	C	Class
F	F	F	Y
F	F	T	N
F	T	T	N
F	T	T	N
F	F	T	Y
T	F	T	Y
T	F	T	N
T	F	T	N
T	T	T	Y
T	F	T	Y

Select the correct prediction class for a test sample (A=F, B=T, C=F) using the naïve Bayes approach.

 'Y'

'N' Anyone**Question 20****2 / 2 pts**

Review the table below.

Instance	X-Axis	Class
i	-3	-
ii	-2	-
iii	-1	+
iv	0	+
v	1	+
vi	2	-

Select a correct hyper-plane equation of the data below using Non-linear SVM.

**Feature Map ( $\{R^1 \rightarrow R^2\}$ )** Y=1 Y=1.5 Y=2 Y=2.5 None of above

Quiz Score: **38** out of 40

# Final Exam

**Due** Dec 8 at 1:15am      **Points** 40      **Questions** 20

**Available** Dec 6 at 12am - Dec 8 at 2am 2 days

**Time Limit** 75 Minutes

## Instructions

No collaboration

75 mins

## Attempt History

	<b>Attempt</b>	<b>Time</b>	<b>Score</b>
<b>LATEST</b>	<a href="#"><u>Attempt 1</u></a>	50 minutes	38 out of 40

! Correct answers are hidden.

Score for this quiz: **38** out of 40

Submitted Dec 6 at 2:37pm

This attempt took 50 minutes.

### Question 1

2 / 2 pts

Example of market basket transactions.

Transaction ID Items Bought

1 {a, b, d, e}

2 {b, c, d}

3 {a, b, d, e}

- 4      {a, c, d, e}
- 5      {b, c, d, e}
- 6      {b, d, e}
- 7      {c, d}
- 8      {a, b, c}
- 9      {a, d, e}
- 10     {b, d}

Compute the support for itemset {b,c,d}.

0.2

## Question 2

2 / 2 pts

Example of market basket transactions.

### Transaction ID Items Bought

- 1      {a, b, d, e}
- 2      {b, c, d}
- 3      {a, b, d, e}
- 4      {a, c, d, e}

5      {b, c, d, e}

6      {b, d, e}

7      {c, d}

8      {a, b, c}

9      {a, d, e}

10     {b, d}

Compute the confidence for the association rules  $\{a, e\} \rightarrow \{b\}$ .

0.5

### Question 3

2 / 2 pts

Transaction ID Items Bought

1      {Milk, Beer, Diapers}

2      {Bread, Butter, Milk}

3      {Milk, Diapers, Cookies}

4      {Bread, Butter, Cookies}

5      {Beer, Cookies, Diapers}

6      {Milk, Diapers, Bread, Butter}

- 7 {Bread, Butter, Diapers}
- 8 {Beer, Diapers}
- 9 {Milk, Diapers, Bread, Butter}
- 10 {Beer, Cookies}

What is the confidence of the rule Beer  $\rightarrow$  Cookies

0.5

#### Question 4

2 / 2 pts

Example of market basket transactions.

#### Transaction ID Items Bought

- 1 {a, b, d, e}
- 2 {b, c, d}
- 3 {a, b, d, e}
- 4 {a, c, d, e}
- 5 {b, c, d, e}
- 6 {b, d, e}
- 7 {c, d}

8            {a, b, c}

9            {a, d, e}

10          {b, d}

If the minimum support is set at 40%, how many frequent 3-itemsets will be found? Note that if a 3 itemset is a subset of a larger itemset, it counts as one occurrence.

2

### Question 5

2 / 2 pts

Transaction ID Items Bought

1            {Milk, Beer, Diapers}

2            {Bread, Butter, Milk}

3            {Milk, Diapers, Cookies}

4            {Bread, Butter, Cookies}

5            {Beer, Cookies, Diapers}

6            {Milk, Diapers, Bread, Butter}

7            {Bread, Butter, Diapers}

8            {Beer, Diapers}

9 {Milk, Diapers, Bread, Butter}

10 {Beer, Cookies}

What is the maximum number of size 3-itemsets that can be derived from this data set. Note that the itemsets derived may not be in the dataset.

20

35

40

15

None of the above.

### Question 6

2 / 2 pts

Which of the following is an application of autoencoders?

feature learning.

dimensionality reduction.

Rule mining

all of the above.

### Question 7

2 / 2 pts

The number of neurons in the output layer should match the number of classes (Where the number of classes is greater than 2) in a supervised learning task.

---

True

---

False

**Question 8****2 / 2 pts**

What is an autoencoder?

---

a neural network that copies its input to its output.

---

a neural network that codes itself.

---

a neural network that maps an output to an input through a hidden layer.

---

a neural network that is trained to attempt to copy its input to its output.

**Question 9****2 / 2 pts**

Consider a movie recommendation application, where an user can log in and select movies they like based on their previous preferences.

Which neural network architecture would be suitable to complete this task

---

Fully-Connected Neural Network.

---

Convolutional Neural Network.

---

- Recurrent Neural Network.
- 
- Restricted Boltzmann Machine.

Incorrect

**Question 10****0 / 2 pts**

Increase in size of a convolutional hidden layer would not necessarily increase the performance of a convolutional network.

- 
- True
- 
- False

**Question 11****2 / 2 pts**

Select different aspects of cluster validation.

- 
- Determining the clustering tendency of a set of data
- 
- To improve the complexity of the algorithm
- 
- Determining the 'correct' number of clusters
- 
- None of the above

correct these are some of the aspects of cluster validation!

**Question 12****2 / 2 pts**

What is the purpose of cluster analysis?

- To avoid finding patterns in noise
- To compare clustering algorithms
- To compare two sets of clusters
- all of the above
- None of the above

we do cluster validity to avoid clustering noise and find a suitable algorithm for our data

**Question 13****2 / 2 pts**

Select qualities of clusters produced by a good clustering algorithm.

- Intra-cluster distances are minimized
- Inter-cluster distances are maximized
- Number of clusters produced
- A and B are correct

good cluster have strong cohesion within the cluster and maximum distance between them.

**Question 14****2 / 2 pts**

Select all statements that are true for Hierarchical clustering.

Although it requires high storage, high computing is not required

Robust to noise

Robust to outliers

It is not easy to decide on the number of cluster because the complicated dendrogram disturbs this decision

None of above

**Question 15****2 / 2 pts**

Review the following table.

Data	X	Y
A	1	2
B	2	2
C	2	3
D	8	7
E	8	8
F	8	2

Select the correct number of core points and DBSCAN clustering result with the parameters below.

Eps: 5

Minimum samples (MinPts):3

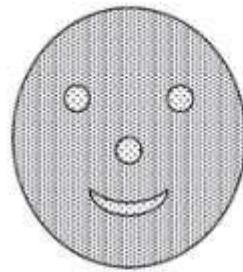
### Distance measurement: Euclidean distance

- 5, cluster1: {A,B,C}, cluster2: {D,E,F}
- 6, cluster1: {A,B}, cluster2: {C,D,E}, cluster3: {F}
- 4, cluster1: {A,B,C}, cluster2: {D, E, F}
- 5, cluster1: {A}, cluster2: {B}, cluster3: {C}, cluster4: {D}, cluster5: {E}
- None of above

### Question 16

2 / 2 pts

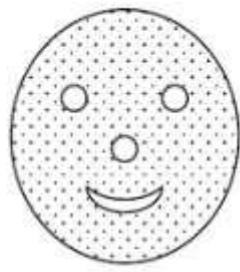
Review the image below.



(a)



(b)



(c)



(d)

**Darkness or number of dots represent density. Lines are used only to distinguish regions and do not represent points.**

If you want to find the patterns represented by the nose, eyes, and mouth using k-Means clustering, select all figures that be well-clustered?

 (a) (b) (c)

(d) None of above**Question 17****2 / 2 pts**

Select all that are not true about DBSCAN.

- it has trouble when the clusters have widely varying densities
- it is very sensitive to noise and to the size of the data
- DBSCAN has a complexity of order  $O(n^3)$
- 
- it can be expensive when the computation of nearest neighbors requires computing all pairwise proximities
- A and B are correct
- B and C are correct

DBSCAN is relatively resistant to noise and can handle clusters of arbitrary shapes and sizes. It has complexity order  $O(n \log n)$

- A, C and D are correct

**Question 18****2 / 2 pts**

Review the table below.

Instance	X-Axis	Class
i	-3	-
ii	-2	-
iii	-1	+
iv	0	+
v	1	+
vi	2	-

Select a correct hyper-plane equation of the data below using Non-linear SVM.

## Feature Map ( $\{R^1 \rightarrow R^2\}$ )

Y=1

Y=1.5

Y=2

Y=2.5

None of above

### Question 19

2 / 2 pts

Review the table below.

Instance	X-Axis	Class
i	-3	-
ii	-2	-
iii	-1	+
iv	0	+
v	1	+
vi	2	-

Select all the support vector instances using Non-linear SVM

## Feature Map ( $\{R^1 \rightarrow R^2\}$ )

i

ii

iii

iv

v

vi

### Question 20

2 / 2 pts

Review the table below.

A	B	C	Class
F	F	F	Y
F	F	T	N
F	T	T	N
F	T	T	N
F	F	T	Y
T	F	T	Y
T	F	T	N
T	F	T	N
T	T	T	Y
T	F	T	Y

Select the correct prediction class for a test sample (A=F, B=T, C=F) using the naïve Bayes approach.

'Y'

'N'

Anyone

Quiz Score: **38** out of 40

# Final Exam

**Due** Dec 9 at 1:15am      **Points** 40      **Questions** 20

**Available** Dec 6 at 12am - Dec 9 at 2am 3 days

**Time Limit** 75 Minutes

## Instructions

No collaboration

75 mins

## Attempt History

	<b>Attempt</b>	<b>Time</b>	<b>Score</b>
LATEST	<a href="#"><u>Attempt 1</u></a>	44 minutes	40 out of 40

! Correct answers are hidden.

Score for this quiz: **40** out of 40

Submitted Dec 8 at 6pm

This attempt took 44 minutes.

### Question 1

2 / 2 pts

Transaction ID Items Bought

- 1 {Milk, Beer, Diapers}
- 2 {Bread, Butter, Milk}
- 3 {Milk, Diapers, Cookies}
- 4 {Bread, Butter, Cookies}

- 5        {Beer, Cookies, Diapers}
- 6        {Milk, Diapers, Bread, Butter}
- 7        {Bread, Butter, Diapers}
- 8        {Beer, Diapers}
- 9        {Milk, Diapers, Bread, Butter}
- 10      {Beer, Cookies}

What is the maximum size of frequent itemsets that can be extracted  
(assuming minsup > 0.1)?

---

10

---

9

---

4

---

3

## Question 2

2 / 2 pts

Which of the above statements are true for any A, B, and C?

---

If A → B then B → A.

---

If A → B and B → C then A → C.

---

If A → C then A union B → C.

---

If A union B  $\rightarrow$  C then A  $\rightarrow$  C.

I & II

I, II, & III

I, II, & IV

II & III

none

### Question 3

2 / 2 pts

Example of market basket transactions.

#### Transaction ID Items Bought

1 {a, b, d, e}

2 {b, c, d}

3 {a, b, d, e}

4 {a, c, d, e}

5 {b, c, d, e}

6 {b, d, e}

7 {c, d}

8 {a, b, c}

9      {a, d, e}

10     {b, d}

Compute the confidence for the association rules  $\{a, e\} \rightarrow \{b\}$ .

0.5

#### Question 4

2 / 2 pts

Transaction ID Items Bought

1      {Milk, Beer, Diapers}

2      {Bread, Butter, Milk}

3      {Milk, Diapers, Cookies}

4      {Bread, Butter, Cookies}

5      {Beer, Cookies, Diapers}

6      {Milk, Diapers, Bread, Butter}

7      {Bread, Butter, Diapers}

8      {Beer, Diapers}

9      {Milk, Diapers, Bread, Butter}

10     {Beer, Cookies}

What is the maximum number of size 3-itemsets that can be derived from this data set. Note that the itemsets derived may not be in the dataset.

20

35

40

15

None of the above.

### Question 5

2 / 2 pts

Transaction ID Items Bought

1 {Milk, Beer, Diapers}

2 {Bread, Butter, Milk}

3 {Milk, Diapers, Cookies}

4 {Bread, Butter, Cookies}

5 {Beer, Cookies, Diapers}

6 {Milk, Diapers, Bread, Butter}

7 {Bread, Butter, Diapers}

8 {Beer, Diapers}

9 {Milk, Diapers, Bread, Butter}

10 {Beer, Cookies}

What is the confidence of the rule Beer -> Cookies

0.5

### Question 6

2 / 2 pts

Increase in size of a convolutional hidden layer would not necessarily increase the performance of a convolutional network.

True

False

### Question 7

2 / 2 pts

Consider a movie recommendation application, where an user can log in and select movies they like based on their previous preferences.

Which neural network architecture would be suitable to complete this task

Fully-Connected Neural Network.

Convolutional Neural Network.

Recurrent Neural Network.

Restricted Boltzmann Machine.

**Question 8****2 / 2 pts**

What is an autoencoder?

- a neural network that copies its input to its output.
- a neural network that codes itself.
- 
- a neural network that maps an output to an input through a hidden layer.
- a neural network that is trained to attempt to copy its input to its output.

**Question 9****2 / 2 pts**

Which of the following is an application of autoencoders?

- feature learning.
- dimensionality reduction.
- Rule mining
- all of the above.

**Question 10****2 / 2 pts**

Is an autoencoder the same as performing Principal Component Analysis?

- Not at all
- No but conceptually they can be used for the same purpose
- Autoencoders can do dimensionality reduction, but it is non-linear.
- Yes they have the same purpose.

**Question 11****2 / 2 pts**

Review the table below.

Cluster	Tennis	Golf	Baseball	Basketball	Football	Hockey	Total	Entropy	Purity
#1	1	1	0	11	4	676	693		
#2	27	89	333	827	253	33	1562		
#3	326	465	8	105	16	29	949		
Total	354	555	341	943	273	738	3204		

Select the collect entropy set of the confusion matrix.

- Cluster #1: 1.84, Cluster #2: 0.2, Cluster #3: 1.7, Total: 1.44
- Cluster #1: 1.44, Cluster #2: 1.84, Cluster #3: 1.7, Total: 0.2
- Cluster #1: 0.2, Cluster #2: 1.7, Cluster #3: 1.84, Total: 1.44
- Cluster #1: 0.2, Cluster #2: 1.44, Cluster #3: 1.7, Total: 1.84
- None of above

Cluster #1: 0.2, Cluster #2: 1.84, Cluster #3: 1.7, Total: 1.44

**Question 12****2 / 2 pts**

Review the table below.

Cluster	Tennis	Golf	Baseball	Basketball	Football	Hockey	Total	Entropy	Purity
#1	1	1	0	11	4	676	693		
#2	27	89	333	827	253	33	1562		
#3	326	465	8	105	16	29	949		
Total	354	555	341	943	273	738	3204		

Select the collect purity set of the confusion matrix

Cluster #1: 0.98, Cluster #2: 0.53, Cluster #3: 0.49, Total: 0.61

Cluster #1: 0.53, Cluster #2: 0.98, Cluster #3: 0.61, Total: 0.49

Cluster #1: 0.98, Cluster #2: 0.49, Cluster #3: 0.53, Total: 0.61

Cluster #1: 0.53, Cluster #2: 0.49, Cluster #3: 0.61, Total: 0.98

None of above

**Question 13****2 / 2 pts**

Review the following table.

	A	B	C	D	E
A	0	9	3	6	11
B	9	0	7	5	10
C	3	7	0	9	2
D	6	5	9	0	8
E	11	10	2	8	0

Select the correct DBSCAN clustering result with the parameters below

epsilon = 5, Minpts = 2

- cluster1: {A, C, E}, cluster2: {B,D}
- cluster1: {A}, cluster2: {B}, cluster3: {D}, and cluster4: {C, E}
- cluster1: {A, B}, cluster2: {C}, and cluster3: {D}, cluster4: {E}
- cluster1: {A}, cluster2: {B, E}, cluster3: {C}, and cluster4: {D}
- None of above

### Question 14

2 / 2 pts

Select qualities of clusters produced by a good clustering algorithm.

- Intra-cluster distances are minimized
- Inter-cluster distances are maximized
- Number of clusters produced
- A and B are correct

good cluster have strong cohesion within the cluster and maximum distance between them.

### Question 15

2 / 2 pts

What type of clustering is DBSCAN?

Center-based Contiguous Well-separated None of above

DBSCAN is a density-based clustering algorithm

**Question 16****2 / 2 pts**

Review the following table.

Data	X	Y
x1	1	1
x2	6	6
x3	7	7
x4	3	2
x5	6	8
x6	2	1
x7	2	3

Compute the average silhouette score to evaluate K-mean clustering result with the parameters below.

K:2

Initial centroids: centroid1 = (4,3) and centroid2 = (9,7)

Distance measure: Euclidean distance

0.752

**Question 17****2 / 2 pts**

Select all that are not true about DBSCAN.

- it has trouble when the clusters have widely varying densities
- it is very sensitive to noise and to the size of the data
- DBSCAN has a complexity of order  $O(n^3)$
- it can be expensive when the computation of nearest neighbors requires computing all pairwise proximities
- A and B are correct
- B and C are correct
  - DBSCAN is relatively resistant to noise and can handle clusters of arbitrary shapes and sizes. It has complexity order  $O(n \log n)$
- A, C and D are correct

**Question 18****2 / 2 pts**

Select the true statement for Nearest Neighbor classification



Although Nearest Neighbor classification has the data dimensionality issue, it can be solved by the scaling technique



k-NN classifier is a typical lazy learner because it spend so much time building the model even if the 'k' is very small



Drawing a circle is a mandatory task to find neighbor classes regardless of the number of 'K'



Determining the optimal value of 'K' is important for k-NN classifier performance.



None of above

### Question 19

2 / 2 pts

Review the table below.

Instance	X-Axis	Class
i	-3	-
ii	-2	-
iii	-1	+
iv	0	+
v	1	+
vi	2	-

Select all the support vector instances using Non-linear SVM

## Feature Map ( $\{R^1 \rightarrow R^2\}$ )

 i ii iii iv v vi

### Question 20

2 / 2 pts

Review the table below.

Instance	X-Axis	Class
i	-3	-
ii	-2	-
iii	-1	+
iv	0	+
v	1	+
vi	2	-

Select a correct hyper-plane equation of the data below using Non-linear SVM.

## Feature Map ( $\{R^1 \rightarrow R^2\}$ )

- Y=1
- Y=1.5
- Y=2
- Y=2.5
- None of above

Quiz Score: **40** out of 40

# Final Exam

**Due** Dec 9 at 1:15am      **Points** 40      **Questions** 20

**Available** Dec 6 at 12am - Dec 9 at 2am 3 days

**Time Limit** 75 Minutes

## Instructions

No collaboration

75 mins

## Attempt History

	Attempt	Time	Score
LATEST	<a href="#"><u>Attempt 1</u></a>	46 minutes	38 out of 40

! Correct answers are hidden.

Score for this quiz: **38** out of 40

Submitted Dec 8 at 4:43pm

This attempt took 46 minutes.

Question 1	2 / 2 pts
Example of market basket transactions.	
Transaction ID Items Bought	
1	{a, b, d, e}
2	{b, c, d}

- |    |              |
|----|--------------|
| 3  | {a, b, d, e} |
| 4  | {a, c, d, e} |
| 5  | {b, c, d, e} |
| 6  | {b, d, e}    |
| 7  | {c, d}       |
| 8  | {a, b, c}    |
| 9  | {a, d, e}    |
| 10 | {b, d}       |

If the minimum support is set at 40%, how many frequent 3-itemsets will be found? Note that if a 3 itemset is a subset of a larger itemset, it counts as one occurrence.

2

**Question 2****2 / 2 pts**

Select the reason why Apriori pruning in the search for frequent itemsets works.

- support count is monotonic with respect to itemsets.
- we search in transaction ID order.
- support count diverges as we add to the itemset.

- support count is anti-monotonic with respect to itemsets.
- it is an excellent heuristic, but it does not work 100% of the time.

**Question 3****2 / 2 pts**

Transaction ID Items Bought

- 1 {Milk, Beer, Diapers}
- 2 {Bread, Butter, Milk}
- 3 {Milk, Diapers, Cookies}
- 4 {Bread, Butter, Cookies}
- 5 {Beer, Cookies, Diapers}
- 6 {Milk, Diapers, Bread, Butter}
- 7 {Bread, Butter, Diapers}
- 8 {Beer, Diapers}
- 9 {Milk, Diapers, Bread, Butter}
- 10 {Beer, Cookies}

What is the maximum size of frequent itemsets that can be extracted  
(assuming minsup > 0.1)?

10 9 4 3**Question 4****2 / 2 pts**

Transaction ID Items Bought

- 1 {Milk, Beer, Diapers}
- 2 {Bread, Butter, Milk}
- 3 {Milk, Diapers, Cookies}
- 4 {Bread, Butter, Cookies}
- 5 {Beer, Cookies, Diapers}
- 6 {Milk, Diapers, Bread, Butter}
- 7 {Bread, Butter, Diapers}
- 8 {Beer, Diapers}
- 9 {Milk, Diapers, Bread, Butter}
- 10 {Beer, Cookies}

Find all itemsets (of size 3 or larger) that has the largest support.

{Milk, Beer, Diapers}

{Bread, Butter, Milk}

{Bread, Butter, Diapers}

{Milk, Diapers, Bread}

None of the above.

All of the above.

**Question 5****2 / 2 pts**

Which of the above statements are true for any A, B, and C?

- If  $A \rightarrow B$  then  $B \rightarrow A$ .
- If  $A \rightarrow B$  and  $B \rightarrow C$  then  $A \rightarrow C$ .
- If  $A \rightarrow C$  then  $A \cup B \rightarrow C$ .
- If  $A \cup B \rightarrow C$  then  $A \rightarrow C$ .
- I & II
- I, II, & III
- I, II, & IV
- II & III
- none

**Question 6****2 / 2 pts**

The number of neurons in the output layer should match the number of classes (Where the number of classes is greater than 2) in a supervised learning task.

- True
- False

**Question 7****2 / 2 pts**

Increase in size of a convolutional hidden layer would not necessarily increase the performance of a convolutional network.

True

False

### Question 8

2 / 2 pts

Is an autoencoder the same as performing Principal Component Analysis?

Not at all

No but conceptually they can be used for the same purpose

Autoencoders can do dimensionality reduction, but it is non-linear.

Yes they have the same purpose.

### Question 9

2 / 2 pts

Usage of deep learning is predicated with availability of large amounts of data. Is the data volume the only determining factor to check whether we can use deep learning ?

Yes, lower volumes lead to overfitting

No, balance among classes in the input data is the only determining factor volume does not matter



No, balance among classes in the input data is a co-factor along with the data volume



None of the above

### Question 10

2 / 2 pts

Which of the following is an application of autoencoders?

feature learning.

dimensionality reduction.

Rule mining

all of the above.

### Question 11

2 / 2 pts

Select all statements that are true for Hierarchical clustering.

Although it requires high storage, high computing is not required

Robust to noise

Robust to outliers

It is not easy to decide on the number of cluster because the complicated dendrogram disturbs this decision

- None of above

### Question 12

2 / 2 pts

Review the table below.

Cluster	Tennis	Golf	Baseball	Basketball	Football	Hockey	Total	Entropy	Purity
#1	1	1	0	11	4	676	693		
#2	27	89	333	827	253	33	1562		
#3	326	465	8	105	16	29	949		
Total	354	555	341	943	273	738	3204		

Select the collect purity set of the confusion matrix

- Cluster #1: 0.98, Cluster #2: 0.53, Cluster #3: 0.49, Total: 0.61

- Cluster #1: 0.53, Cluster #2: 0.98, Cluster #3: 0.61, Total: 0.49

- Cluster #1: 0.98, Cluster #2: 0.49, Cluster #3: 0.53, Total: 0.61

- Cluster #1: 0.53, Cluster #2: 0.49, Cluster #3: 0.61, Total: 0.98

- None of above

### Question 13

2 / 2 pts

Select all that are not true about DBSCAN.

- It has trouble when the clusters have widely varying densities

- it is very sensitive to noise and to the size of the data
- DBSCAN has a complexity of order  $O(n^3)$
- it can be expensive when the computation of nearest neighbors requires computing all pairwise proximities
- A and B are correct
- B and C are correct

DBSCAN is relatively resistant to noise and can handle clusters of arbitrary shapes and sizes. It has complexity order  $O(n \log n)$

- A, C and D are correct

### Question 14

2 / 2 pts

Review the following table.

Data	X	Y
x1	1	1
x2	6	6
x3	7	7
x4	3	2
x5	6	8
x6	2	1
x7	2	3

Select the correct pair of final centroid and clustering set using K-mean clustering with the parameters below.

K:2

Initial centroids: centroid1 = (4,3) and centroid2 = (9,7)

Distance measure: Euclidean distance



Cluster 1: (x1, x4, x6, x7) with centroid=(2, 1.75), Cluster 2: (x2, x3, x5) with centroid=(6.33, 7)



Cluster 1: (x1, x4, x6, x7) with centroid=(1.75, 2), Cluster 2: (x2, x3, x5) with centroid=(6.33, 7)



Cluster 1: (x1, x4, x6, x7) with centroid=(2, 1.75), Cluster 2: (x2, x3, x5) with centroid=(7, 6.33)



Cluster 1: (x1, x2, x4, x6, x7) with centroid=(2, 1.75), Cluster 2: (x3, x5) with centroid=(6.33, 7)



Cluster 1: (x1, x2, x4, x6, x7) with centroid=(2, 1.75), Cluster 2: (x3, x5) with centroid=(7, 6.33)

### Question 15

2 / 2 pts

Review the following table.

Data	X	Y
x1	1	1
x2	6	6
x3	7	7
x4	3	2
x5	6	8
x6	2	1
x7	2	3

Compute the sum of squared error (SSE) to evaluate K-mean clustering result with the parameters below.

K:2

Initial centroids: centroid1 = (4,3) and centroid2 = (9,7)

Distance measure: Euclidean distance

7.4167

Incorrect

### Question 16

0 / 2 pts

Review the following table.

Data	X	Y
x1	1	1
x2	6	6
x3	7	7
x4	3	2
x5	6	8
x6	2	1
x7	2	3

Compute the updated centroids after the first iteration using K-mean clustering with the parameters below.

K:2

Initial centroids: centroid1 = (4,3) and centroid2 = (9,7)

Distance measure: Euclidean

Type your answer in the same format as "centroid1 = (x1,y1) and centroid2 = (x2,y2)"

Each co-ordinate should be in the format A.BC, where A, B and C are integers. If you have 4 as one of the co-ordinates you should write 4.00. Please be mindful of the spaces and other formats.

centroid1 = (2.00,1.75) and centroid2 = (6.33,7.00)

### Question 17

2 / 2 pts

Review the following table.

	A	B	C	D	E
A	0	9	3	6	11
B	9	0	7	5	10
C	3	7	0	9	2
D	6	5	9	0	8
E	11	10	2	8	0

Select the correct DBSCAN clustering result with the parameters below

epsilon = 5, Minpts = 2

- cluster1: {A, C, E}, cluster2: {B,D}
- cluster1: {A}, cluster2: {B}, cluster3: {D}, and cluster4: {C, E}
- cluster1: {A, B}, cluster2: {C}, and cluster3: {D}, cluster4: {E}
- cluster1: {A}, cluster2: {B, E}, cluster3: {C}, and cluster4: {D}
- None of above

### Question 18

2 / 2 pts

Review the table below.

A	B	C	Class
F	F	F	Y
F	F	T	N
F	T	T	N
F	T	T	N
F	F	T	Y
T	F	T	Y
T	F	T	N
T	F	T	N
T	T	T	Y
T	F	T	Y

Select the correct prediction class for a test sample (A=F, B=T, C=F) using the naïve Bayes approach.

'Y'

'N'

Anyone

### Question 19

2 / 2 pts

Select the true statement for Nearest Neighbor classification

Although Nearest Neighbor classification has the data dimensionality issue, it can be solved by the scaling technique

k-NN classifier is a typical lazy learner because it spend so much time building the model even if the 'k' is very small

Drawing a circle is a mandatory task to find neighbor classes regardless of the number of 'K'

Determining the optimal value of 'K' is important for k-NN classifier performance.

None of above

### Question 20

2 / 2 pts

Review the table below.

Instance	X-Axis	Class
i	-3	-
ii	-2	-
iii	-1	+
iv	0	+
v	1	+
vi	2	-

Select all the support vector instances using Non-linear SVM

Feature Map ( $\{R^1 \rightarrow R^2\}$ )

 i

<input checked="" type="checkbox"/> ii
<input checked="" type="checkbox"/> iii
<input type="checkbox"/> iv
<input checked="" type="checkbox"/> v
<input checked="" type="checkbox"/> vi

Quiz Score: **38** out of 40

# Final Exam

**Due** Dec 9 at 1:15am      **Points** 40      **Questions** 20

**Available** Dec 6 at 12am - Dec 9 at 2am 3 days

**Time Limit** 75 Minutes

## Instructions

No collaboration

75 mins

## Attempt History

	<b>Attempt</b>	<b>Time</b>	<b>Score</b>
LATEST	<a href="#">Attempt 1</a>	43 minutes	38 out of 40

! Correct answers are hidden.

Score for this quiz: 38 out of 40

Submitted Dec 8 at 5:27pm

This attempt took 43 minutes.

### Question 1

2 / 2 pts

Transaction ID Items Bought

- 1 {Milk, Beer, Diapers}
- 2 {Bread, Butter, Milk}
- 3 {Milk, Diapers, Cookies}
- 4 {Bread, Butter, Cookies}

- 5 {Beer, Cookies, Diapers}
- 6 {Milk, Diapers, Bread, Butter}
- 7 {Bread, Butter, Diapers}
- 8 {Beer, Diapers}
- 9 {Milk, Diapers, Bread, Butter}
- 10 {Beer, Cookies}

What is the confidence of the rule Beer  $\rightarrow$  Cookies

0.5

## Question 2

2 / 2 pts

Transaction ID Items Bought

- 1 {Milk, Beer, Diapers}
- 2 {Bread, Butter, Milk}
- 3 {Milk, Diapers, Cookies}
- 4 {Bread, Butter, Cookies}
- 5 {Beer, Cookies, Diapers}
- 6 {Milk, Diapers, Bread, Butter}

- 7      {Bread, Butter, Diapers}
- 8      {Beer, Diapers}
- 9      {Milk, Diapers, Bread, Butter}
- 10     {Beer, Cookies}

What is the maximum number of size 3-itemsets that can be derived from this data set. Note that the itemsets derived may not be in the dataset.

- 
- 20
- 
- 35
- 
- 40
- 
- 15
- 
- None of the above.

### Question 3

2 / 2 pts

Example of market basket transactions.

Transaction ID Items Bought

- 1      {a, b, d, e}
- 2      {b, c, d}
- 3      {a, b, d, e}

- 4      {a, c, d, e}
- 5      {b, c, d, e}
- 6      {b, d, e}
- 7      {c, d}
- 8      {a, b, c}
- 9      {a, d, e}
- 10     {b, d}

If the minimum support is set at 40%, how many frequent 3-itemsets will be found? Note that if a 3 itemset is a subset of a larger itemset, it counts as one occurrence.

2

**Question 4****2 / 2 pts**

Transaction ID Items Bought

- 1      {Milk, Beer, Diapers}
- 2      {Bread, Butter, Milk}
- 3      {Milk, Diapers, Cookies}
- 4      {Bread, Butter, Cookies}

- 5      {Beer, Cookies, Diapers}
- 6      {Milk, Diapers, Bread, Butter}
- 7      {Bread, Butter, Diapers}
- 8      {Beer, Diapers}
- 9      {Milk, Diapers, Bread, Butter}
- 10     {Beer, Cookies}

Find all itemsets (of size 3 or larger) that has the largest support.

- 
- {Milk, Beer, Diapers}
- 
- {Bread, Butter, Milk}
- 
- {Bread, Butter, Diapers}
- 
- {Milk, Diapers, Bread}
- 
- None of the above.
- 
- All of the above.

### Question 5

2 / 2 pts

Example of market basket transactions.

Transaction ID Items Bought

- 1      {a, b, d, e}

- 2      {b, c, d}
- 3      {a, b, d, e}
- 4      {a, c, d, e}
- 5      {b, c, d, e}
- 6      {b, d, e}
- 7      {c, d}
- 8      {a, b, c}
- 9      {a, d, e}
- 10     {b, d}

Compute the confidence for the association rules  $\{a, e\} \rightarrow \{b\}$ .

0.5

## Question 6

2 / 2 pts

Usage of deep learning is predicated with availability of large amounts of data. Is the data volume the only determining factor to check whether we can use deep learning ?

Yes, lower volumes lead to overfitting



No, balance among classes in the input data is the only determining factor  
volume does not matter



No, balance among classes in the input data is a co-factor along with the data volume



None of the above

### Question 7

2 / 2 pts

Select all that is true about Restricted Boltzman Machine (RBM)

Can be used for recommendation systems

RBM training is often probabilistic

RBM is a kind of Deep Belief Network

None of the above.

### Question 8

2 / 2 pts

Which of the following is an application of autoencoders?

feature learning.

dimensionality reduction.

Rule mining

all of the above.

**Question 9****2 / 2 pts**

Increase in size of a convolutional hidden layer would not necessarily increase the performance of a convolutional network.

---

True

---

False

**Question 10****2 / 2 pts**

Select the true statements about neural network (NN) complexity.

---

CNN is always more complex than DNN

---

Adding more hidden layers increases complexity of a NN

---

An Autoencoder is always more complex than a CNN

---



Complexity metric defined for a NN may not make sense for other classification engines

**Question 11****2 / 2 pts**

Review the following table.

Data	X	Y
A	1	2
B	2	2

C	2	3
D	8	7
E	8	8
F	8	2

Select the correct number of core points and DBSCAN clustering result with the parameters below.

Eps: 5

Minimum samples (MinPts):3

Distance measurement: Euclidean distance

- 5, cluster1: {A,B,C}, cluster2: {D,E,F}
- 6, cluster1: {A,B}, cluster2: {C,D,E}, cluster3: {F}
- 4, cluster1: {A,B,C}, cluster2: {D, E, F}
- 5, cluster1: {A}, cluster2: {B}, cluster3: {C}, cluster4: {D}, cluster5: {E}
- None of above

### Question 12

2 / 2 pts

Select qualities of clusters produced by a good clustering algorithm.

- Intra-cluster distances are minimized
- Inter-cluster distances are maximized
- Number of clusters produced
- A and B are correct

good cluster have strong cohesion within the cluster and maximum distance between them.

**Question 13****2 / 2 pts**

What is the purpose of cluster analysis?

- To avoid finding patterns in noise
- To compare clustering algorithms
- To compare two sets of clusters
- all of the above
- None of the above

we do cluster validity to avoid clustering noise and find a suitable algorithim for our data

**Question 14****2 / 2 pts**

Review the following table.

Data	X	Y
x1	1	1
x2	6	6
x3	7	7
x4	3	2
x5	6	8
x6	2	1
x7	2	3

Select the correct pair of final centroid and clustering set using K-mean clustering with the parameters below.

K:2

Initial centroids: centroid1 = (4,3) and centroid2 = (9,7)

Distance measure: Euclidean distance



Cluster 1: (x1, x4, x6, x7) with centroid=(2, 1.75), Cluster 2: (x2, x3, x5) with centroid=(6.33, 7)



Cluster 1: (x1, x4, x6, x7) with centroid=(1.75, 2), Cluster 2: (x2, x3, x5) with centroid=(6.33, 7)



Cluster 1: (x1, x4, x6, x7) with centroid=(2, 1.75), Cluster 2: (x2, x3, x5) with centroid=(7, 6.33)



Cluster 1: (x1, x2, x4, x6, x7) with centroid=(2, 1.75), Cluster 2: (x3, x5) with centroid=(6.33, 7)



Cluster 1:  $(x_1, x_2, x_4, x_6, x_7)$  with centroid=(2, 1.75), Cluster 2:  $(x_3, x_5)$  with centroid=(7, 6.33)

**Question 15****2 / 2 pts**

Review the following table.

Data	X	Y
x1	1	1
x2	6	6
x3	7	7
x4	3	2
x5	6	8
x6	2	1
x7	2	3

Compute the sum of squared error (SSE) to evaluate K-mean clustering result with the parameters below.

K:2

Initial centroids: centroid1 = (4,3) and centroid2 = (9,7)

Distance measure: Euclidean distance

7.4167

**Question 16****2 / 2 pts**

Select all statements that are true for Hierarchical clustering.

Although it requires high storage, high computing is not required

Robust to noise

Robust to outliers

It is not easy to decide on the number of cluster because the complicated dendrogram disturbs this decision

None of above

Incorrect

### Question 17

0 / 2 pts

Review the following table.

Data	X	Y
x1	1	1
x2	6	6
x3	7	7
x4	3	2
x5	6	8
x6	2	1
x7	2	3

Compute the updated centroids after the first iteration using K-mean clustering with the parameters below.

K:2

Initial centroids: centroid1 = (4,3) and centroid2 = (9,7)

Distance measure: Euclidean

Type your answer in the same format as "centroid1 = (x1,y1) and centroid2 = (x2,y2)"

Each co-ordinate should be in the format A.BC, where A, B and C are integers. If you have 4 as one of the co-ordinates you should write 4.00. Please be mindful of the spaces and other formats.

centroid1 = (2.00,1.75) and centroid2 = (6.33,7.00)

### Question 18

2 / 2 pts

Review the table below.

Instance	X-Axis	Class
i	-3	-
ii	-2	-
iii	-1	+
iv	0	+
v	1	+
vi	2	-

Select all the support vector instances using Non-linear SVM

Feature Map ( $\{R^1 \rightarrow R^2\}$ )

i

ii iii iv v vi**Question 19****2 / 2 pts**

Review the table below.

A	B	C	Class
F	F	F	Y
F	F	T	N
F	T	T	N
F	T	T	N
F	F	T	Y
T	F	T	Y
T	F	T	N
T	F	T	N
T	T	T	Y
T	F	T	Y

Select the correct prediction class for a test sample (A=F, B=T, C=F) using the naïve Bayes approach.

 'Y'

'N' Anyone**Question 20****2 / 2 pts**

Review the table below.

Instance	X-Axis	Class
i	-3	-
ii	-2	-
iii	-1	+
iv	0	+
v	1	+
vi	2	-

Select a correct hyper-plane equation of the data below using Non-linear SVM.

**Feature Map ( $\{R^1 \rightarrow R^2\}$ )** Y=1 Y=1.5 Y=2 Y=2.5 None of above

Quiz Score: **38** out of 40

# Final Exam

**Due** Dec 9 at 1:15am      **Points** 40      **Questions** 20

**Available** Dec 6 at 12am - Dec 9 at 2am 3 days

**Time Limit** 75 Minutes

## Instructions

No collaboration

75 mins

## Attempt History

	Attempt	Time	Score
LATEST	<a href="#"><u>Attempt 1</u></a>	46 minutes	38 out of 40

! Correct answers are hidden.

Score for this quiz: **38** out of 40

Submitted Dec 8 at 4:43pm

This attempt took 46 minutes.

Question 1	2 / 2 pts
Example of market basket transactions.	
Transaction ID Items Bought	
1	{a, b, d, e}
2	{b, c, d}

- |    |              |
|----|--------------|
| 3  | {a, b, d, e} |
| 4  | {a, c, d, e} |
| 5  | {b, c, d, e} |
| 6  | {b, d, e}    |
| 7  | {c, d}       |
| 8  | {a, b, c}    |
| 9  | {a, d, e}    |
| 10 | {b, d}       |

If the minimum support is set at 40%, how many frequent 3-itemsets will be found? Note that if a 3 itemset is a subset of a larger itemset, it counts as one occurrence.

2

**Question 2****2 / 2 pts**

Select the reason why Apriori pruning in the search for frequent itemsets works.

- support count is monotonic with respect to itemsets.
- we search in transaction ID order.
- support count diverges as we add to the itemset.

- support count is anti-monotonic with respect to itemsets.
- it is an excellent heuristic, but it does not work 100% of the time.

**Question 3****2 / 2 pts**

Transaction ID Items Bought

- 1 {Milk, Beer, Diapers}
- 2 {Bread, Butter, Milk}
- 3 {Milk, Diapers, Cookies}
- 4 {Bread, Butter, Cookies}
- 5 {Beer, Cookies, Diapers}
- 6 {Milk, Diapers, Bread, Butter}
- 7 {Bread, Butter, Diapers}
- 8 {Beer, Diapers}
- 9 {Milk, Diapers, Bread, Butter}
- 10 {Beer, Cookies}

What is the maximum size of frequent itemsets that can be extracted  
(assuming minsup > 0.1)?

10 9 4 3**Question 4****2 / 2 pts**

Transaction ID Items Bought

- 1 {Milk, Beer, Diapers}
- 2 {Bread, Butter, Milk}
- 3 {Milk, Diapers, Cookies}
- 4 {Bread, Butter, Cookies}
- 5 {Beer, Cookies, Diapers}
- 6 {Milk, Diapers, Bread, Butter}
- 7 {Bread, Butter, Diapers}
- 8 {Beer, Diapers}
- 9 {Milk, Diapers, Bread, Butter}
- 10 {Beer, Cookies}

Find all itemsets (of size 3 or larger) that has the largest support.

{Milk, Beer, Diapers}

{Bread, Butter, Milk}

{Bread, Butter, Diapers}

{Milk, Diapers, Bread}

None of the above.

All of the above.

**Question 5****2 / 2 pts**

Which of the above statements are true for any A, B, and C?

- If  $A \rightarrow B$  then  $B \rightarrow A$ .
- If  $A \rightarrow B$  and  $B \rightarrow C$  then  $A \rightarrow C$ .
- If  $A \rightarrow C$  then  $A \cup B \rightarrow C$ .
- If  $A \cup B \rightarrow C$  then  $A \rightarrow C$ .
- I & II
- I, II, & III
- I, II, & IV
- II & III
- none

**Question 6****2 / 2 pts**

The number of neurons in the output layer should match the number of classes (Where the number of classes is greater than 2) in a supervised learning task.

- True
- False

**Question 7****2 / 2 pts**

Increase in size of a convolutional hidden layer would not necessarily increase the performance of a convolutional network.

True

False

### Question 8

2 / 2 pts

Is an autoencoder the same as performing Principal Component Analysis?

Not at all

No but conceptually they can be used for the same purpose

Autoencoders can do dimensionality reduction, but it is non-linear.

Yes they have the same purpose.

### Question 9

2 / 2 pts

Usage of deep learning is predicated with availability of large amounts of data. Is the data volume the only determining factor to check whether we can use deep learning ?

Yes, lower volumes lead to overfitting

No, balance among classes in the input data is the only determining factor volume does not matter



No, balance among classes in the input data is a co-factor along with the data volume



None of the above

### Question 10

2 / 2 pts

Which of the following is an application of autoencoders?

feature learning.

dimensionality reduction.

Rule mining

all of the above.

### Question 11

2 / 2 pts

Select all statements that are true for Hierarchical clustering.

Although it requires high storage, high computing is not required

Robust to noise

Robust to outliers

It is not easy to decide on the number of cluster because the complicated dendrogram disturbs this decision

- None of above

### Question 12

2 / 2 pts

Review the table below.

Cluster	Tennis	Golf	Baseball	Basketball	Football	Hockey	Total	Entropy	Purity
#1	1	1	0	11	4	676	693		
#2	27	89	333	827	253	33	1562		
#3	326	465	8	105	16	29	949		
Total	354	555	341	943	273	738	3204		

Select the collect purity set of the confusion matrix

- Cluster #1: 0.98, Cluster #2: 0.53, Cluster #3: 0.49, Total: 0.61

- Cluster #1: 0.53, Cluster #2: 0.98, Cluster #3: 0.61, Total: 0.49

- Cluster #1: 0.98, Cluster #2: 0.49, Cluster #3: 0.53, Total: 0.61

- Cluster #1: 0.53, Cluster #2: 0.49, Cluster #3: 0.61, Total: 0.98

- None of above

### Question 13

2 / 2 pts

Select all that are not true about DBSCAN.

- It has trouble when the clusters have widely varying densities

- it is very sensitive to noise and to the size of the data
- DBSCAN has a complexity of order  $O(n^3)$
- it can be expensive when the computation of nearest neighbors requires computing all pairwise proximities
- A and B are correct
- B and C are correct

DBSCAN is relatively resistant to noise and can handle clusters of arbitrary shapes and sizes. It has complexity order  $O(n \log n)$

- A, C and D are correct

### Question 14

2 / 2 pts

Review the following table.

Data	X	Y
x1	1	1
x2	6	6
x3	7	7
x4	3	2
x5	6	8
x6	2	1
x7	2	3

Select the correct pair of final centroid and clustering set using K-mean clustering with the parameters below.

K:2

Initial centroids: centroid1 = (4,3) and centroid2 = (9,7)

Distance measure: Euclidean distance



Cluster 1: (x1, x4, x6, x7) with centroid=(2, 1.75), Cluster 2: (x2, x3, x5) with centroid=(6.33, 7)



Cluster 1: (x1, x4, x6, x7) with centroid=(1.75, 2), Cluster 2: (x2, x3, x5) with centroid=(6.33, 7)



Cluster 1: (x1, x4, x6, x7) with centroid=(2, 1.75), Cluster 2: (x2, x3, x5) with centroid=(7, 6.33)



Cluster 1: (x1, x2, x4, x6, x7) with centroid=(2, 1.75), Cluster 2: (x3, x5) with centroid=(6.33, 7)



Cluster 1: (x1, x2, x4, x6, x7) with centroid=(2, 1.75), Cluster 2: (x3, x5) with centroid=(7, 6.33)

### Question 15

2 / 2 pts

Review the following table.

Data	X	Y
x1	1	1
x2	6	6
x3	7	7
x4	3	2
x5	6	8
x6	2	1
x7	2	3

Compute the sum of squared error (SSE) to evaluate K-mean clustering result with the parameters below.

K:2

Initial centroids: centroid1 = (4,3) and centroid2 = (9,7)

Distance measure: Euclidean distance

7.4167

Incorrect

### Question 16

0 / 2 pts

Review the following table.

Data	X	Y
x1	1	1
x2	6	6
x3	7	7
x4	3	2
x5	6	8
x6	2	1
x7	2	3

Compute the updated centroids after the first iteration using K-mean clustering with the parameters below.

K:2

Initial centroids: centroid1 = (4,3) and centroid2 = (9,7)

Distance measure: Euclidean

Type your answer in the same format as "centroid1 = (x1,y1) and centroid2 = (x2,y2)"

Each co-ordinate should be in the format A.BC, where A, B and C are integers. If you have 4 as one of the co-ordinates you should write 4.00. Please be mindful of the spaces and other formats.

centroid1 = (2.00,1.75) and centroid2 = (6.33,7.00)

### Question 17

2 / 2 pts

Review the following table.

	A	B	C	D	E
A	0	9	3	6	11
B	9	0	7	5	10
C	3	7	0	9	2
D	6	5	9	0	8
E	11	10	2	8	0

Select the correct DBSCAN clustering result with the parameters below

epsilon = 5, Minpts = 2

- cluster1: {A, C, E}, cluster2: {B,D}
- cluster1: {A}, cluster2: {B}, cluster3: {D}, and cluster4: {C, E}
- cluster1: {A, B}, cluster2: {C}, and cluster3: {D}, cluster4: {E}
- cluster1: {A}, cluster2: {B, E}, cluster3: {C}, and cluster4: {D}
- None of above

### Question 18

2 / 2 pts

Review the table below.

A	B	C	Class
F	F	F	Y
F	F	T	N
F	T	T	N
F	T	T	N
F	F	T	Y
T	F	T	Y
T	F	T	N
T	F	T	N
T	T	T	Y
T	F	T	Y

Select the correct prediction class for a test sample (A=F, B=T, C=F) using the naïve Bayes approach.

'Y'

'N'

Anyone

### Question 19

2 / 2 pts

Select the true statement for Nearest Neighbor classification

Although Nearest Neighbor classification has the data dimensionality issue, it can be solved by the scaling technique

k-NN classifier is a typical lazy learner because it spend so much time building the model even if the 'k' is very small

Drawing a circle is a mandatory task to find neighbor classes regardless of the number of 'K'

Determining the optimal value of 'K' is important for k-NN classifier performance.

None of above

### Question 20

2 / 2 pts

Review the table below.

Instance	X-Axis	Class
i	-3	-
ii	-2	-
iii	-1	+
iv	0	+
v	1	+
vi	2	-

Select all the support vector instances using Non-linear SVM

Feature Map ( $\{R^1 \rightarrow R^2\}$ )

 i

<input checked="" type="checkbox"/> ii
<input checked="" type="checkbox"/> iii
<input type="checkbox"/> iv
<input checked="" type="checkbox"/> v
<input checked="" type="checkbox"/> vi

Quiz Score: **38** out of 40

# Final Exam

**Due Dec 9 at 1:15am      Points 40      Questions 20**

**Available** Dec 6 at 12am - Dec 9 at 2am 3 days

**Time Limit** 75 Minutes

## Instructions

No collaboration

75 mins

## Attempt History

	<b>Attempt</b>	<b>Time</b>	<b>Score</b>
<b>LATEST</b>	<a href="#"><u>Attempt 1</u></a>	41 minutes	40 out of 40

! Correct answers are hidden.

Score for this quiz: **40** out of 40

Submitted Dec 8 at 6:17pm

This attempt took 41 minutes.

<b>Question 1</b>	<b>2 / 2 pts</b>
Apriori pruning based on support is a greedy strategy but is not optimal	
<input type="radio"/> True	
<input checked="" type="radio"/> False	

**Question 2****2 / 2 pts**

Which of the above statements are true for any A, B, and C?

- If  $A \rightarrow B$  then  $B \rightarrow A$ .
- If  $A \rightarrow B$  and  $B \rightarrow C$  then  $A \rightarrow C$ .
- If  $A \rightarrow C$  then  $A \cup B \rightarrow C$ .
- If  $A \cup B \rightarrow C$  then  $A \rightarrow C$ .
- I & II
- I, II, & III
- I, II, & IV
- II & III
- none

**Question 3****2 / 2 pts**

Example of market basket transactions.

Transaction ID Items Bought

- |    |              |
|----|--------------|
| 1  | {a, b, d, e} |
| 2  | {b, c, d}    |
| 3  | {a, b, d, e} |
| 4  | {a, c, d, e} |
| 5  | {b, c, d, e} |
| 6  | {b, d, e}    |
| 7  | {c, d}       |
| 8  | {a, b, c}    |
| 9  | {a, d, e}    |
| 10 | {b, d}       |

Compute the confidence for the association rules {a, e} -> {b}.

0.5

#### Question 4

2 / 2 pts

Example of market basket transactions.

**Transaction ID Items Bought**

1 {a, b, d, e}

2 {b, c, d}

3 {a, b, d, e}

4 {a, c, d, e}

5 {b, c, d, e}

6 {b, d, e}

7 {c, d}

8 {a, b, c}

9 {a, d, e}

10 {b, d}

If the minimum support is set at 40%, how many frequent 3-itemsets will be found? Note that if a 3 itemset is a subset of a larger itemset, it counts as one occurrence.

2

**Question 5****2 / 2 pts**

### Example of market basket transactions.

#### Transaction ID Items Bought

1 {a, b, d, e}

2 {b, c, d}

3 {a, b, d, e}

4 {a, c, d, e}

5 {b, c, d, e}

6 {b, d, e}

7 {c, d}

8 {a, b, c}

9 {a, d, e}

10 {b, d}

Compute the support for itemset {b,c,d}.

0.2

**Question 6****2 / 2 pts**

Which of the following is an application of autoencoders?

- 
- feature learning.
- 
- dimensionality reduction.
- 
- Rule mining
- 
- all of the above.

**Question 7****2 / 2 pts**

What is an autoencoder?

- 
- a neural network that copies its input to its output.
- 
- a neural network that codes itself.
- 
- 
- a neural network that maps an output to an input through a hidden layer.
- 
- 
- a neural network that is trained to attempt to copy its input to its output.

**Question 8****2 / 2 pts**

The number of neurons in the output layer should match the number of classes (Where the number of classes is greater than 2) in a supervised learning task.

 True False**Question 9****2 / 2 pts**

Select all that is true about Restricted Boltzman Machine (RBM)

 Can be used for recommendation systems RBM training is often probabilistic RBM is a kind of Deep Belief Network None of the above.**Question 10****2 / 2 pts**

Consider a movie recommendation application, where an user can log in and select movies they like based on their previous preferences.

Which neural network architecture would be suitable to complete this task

- Fully-Connected Neural Network.
- Convolutional Neural Network.
- Recurrent Neural Network.
- Restricted Boltzmann Machine.

### Question 11

2 / 2 pts

Review the following table.

Data	X	Y
x1	1	1
x2	6	6
x3	7	7
x4	3	2
x5	6	8
x6	2	1
x7	2	3

Compute the average silhouette score to evaluate K-mean clustering result with the parameters below.

K:2

Initial centroids: centroid1 = (4,3) and centroid2 = (9,7)

Distance measure: Euclidean distance

0.752

### Question 12

2 / 2 pts

Review the following table.

Data	X	Y
A	1	2
B	2	2
C	2	3
D	8	7
E	8	8
F	8	2

Select the correct number of core points and DBSCAN clustering result with the parameters below.

Eps: 5

Minimum samples (MinPts):3

Distance measurement: Euclidean distance

- 5, cluster1: {A,B,C}, cluster2: {D,E,F}
- 6, cluster1: {A,B}, cluster2: {C,D,E}, cluster3: {F}
- 4, cluster1: {A,B,C}, cluster2: {D, E, F}



5, cluster1: {A}, cluster2: {B}, cluster3: {C}, cluster4: {D}, cluster5: {E}



None of above

### Question 13

2 / 2 pts

Review the following table.

Data	X	Y
x1	1	1
x2	6	6
x3	7	7
x4	3	2
x5	6	8
x6	2	1
x7	2	3

Compute the sum of squared error (SSE) to evaluate K-mean clustering result with the parameters below.

K:2

Initial centroids: centroid1 = (4,3) and centroid2 = (9,7)

Distance measure: Euclidean distance

7.4167

**Question 14****2 / 2 pts**

Select different aspects of cluster validation.

- Determining the clustering tendency of a set of data
- To improve the complexity of the algorithm
- Determining the 'correct' number of clusters
- None of the above

correct these are some of the aspects of cluster validation!

**Question 15****2 / 2 pts**

Select qualities of clusters produced by a good clustering algorithm.

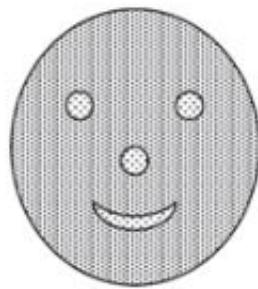
- Intra-cluster distances are minimized
- Inter-cluster distances are maximized
- Number of clusters produced
- A and B are correct

good cluster have strong cohesion within the cluster and maximum distance between them.

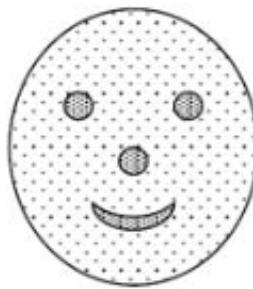
### Question 16

2 / 2 pts

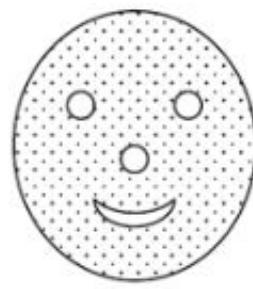
Review the image below.



(a)



(b)



(c)



(d)

**Darkness or number of dots represent density. Lines are used only to distinguish regions and do not represent points.**

If we want to find the patterns represented by the nose, eyes, and mouth using single linkage hierarchical clustering, select all figures that be well-clustered?

(a)

(b)

(c)

(d)

- None of above

**Question 17****2 / 2 pts**

Select all that are not true about DBSCAN.

- it has trouble when the clusters have widely varying densities
- it is very sensitive to noise and to the size of the data
- DBSCAN has a complexity of order  $O(n^3)$
- it can be expensive when the computation of nearest neighbors requires computing all pairwise proximities
- A and B are correct
- B and C are correct

DBSCAN is relatively resistant to noise and can handle clusters of arbitrary shapes and sizes. It has complexity order  $O(n \log n)$
- A, C and D are correct

**Question 18****2 / 2 pts**

### Select the true statement for Nearest Neighbor classification



Although Nearest Neighbor classification has the data dimensionality issue, it can be solved by the scaling technique



k-NN classifier is a typical lazy learner because it spend so much time building the model even if the 'k' is very small



Drawing a circle is a mandatory task to find neighbor classes regardless of the number of 'K'



Determining the optimal value of 'K' is important for k-NN classifier performance.



None of above

### Question 19

2 / 2 pts

Review the table below.

Rainy	Sunshine		Wind		Atmosphere		Temperature	
	Yes	No	Yes	No	High	Low	High	Low
Yes	2	6	6	2	8	0	5	3
No	8	4	2	10	2	10	6	6

Is it raining if the weather condition is *sunshine, no wind, high atmosphere, and low temperature* (Hint: you should use Naïve Bayes)?

Yes

No

## Question 20

2 / 2 pts

Review the table below.

A	B	C	Class
F	F	F	Y
F	F	T	N
F	T	T	N
F	T	T	N
F	F	T	Y
T	F	T	Y
T	F	T	N
T	F	T	N
T	T	T	Y
T	F	T	Y

Select the correct prediction class for a test sample (A=F, B=T, C=F) using the naïve Bayes approach.

---

'Y'

---

'N'

---

Anyone

---

Quiz Score: **40** out of 40

# Final Exam

**Due** Dec 9 at 1:15am      **Points** 40      **Questions** 20

**Available** Dec 6 at 12am - Dec 9 at 2am 3 days

**Time Limit** 75 Minutes

## Instructions

No collaboration

75 mins

## Attempt History

	<b>Attempt</b>	<b>Time</b>	<b>Score</b>
LATEST	<a href="#"><u>Attempt 1</u></a>	44 minutes	40 out of 40

! Correct answers are hidden.

Score for this quiz: **40** out of 40

Submitted Dec 8 at 6pm

This attempt took 44 minutes.

### Question 1

2 / 2 pts

Transaction ID Items Bought

- 1 {Milk, Beer, Diapers}
- 2 {Bread, Butter, Milk}
- 3 {Milk, Diapers, Cookies}
- 4 {Bread, Butter, Cookies}

- 5        {Beer, Cookies, Diapers}
- 6        {Milk, Diapers, Bread, Butter}
- 7        {Bread, Butter, Diapers}
- 8        {Beer, Diapers}
- 9        {Milk, Diapers, Bread, Butter}
- 10      {Beer, Cookies}

What is the maximum size of frequent itemsets that can be extracted  
(assuming minsup > 0.1)?

---

10

---

9

---

4

---

3

## Question 2

2 / 2 pts

Which of the above statements are true for any A, B, and C?

---

If A → B then B → A.

---

If A → B and B → C then A → C.

---

If A → C then A union B → C.

---

If A union B  $\rightarrow$  C then A  $\rightarrow$  C.

I & II

I, II, & III

I, II, & IV

II & III

none

### Question 3

2 / 2 pts

Example of market basket transactions.

#### Transaction ID Items Bought

1 {a, b, d, e}

2 {b, c, d}

3 {a, b, d, e}

4 {a, c, d, e}

5 {b, c, d, e}

6 {b, d, e}

7 {c, d}

8 {a, b, c}

9      {a, d, e}

10     {b, d}

Compute the confidence for the association rules  $\{a, e\} \rightarrow \{b\}$ .

0.5

#### Question 4

2 / 2 pts

Transaction ID Items Bought

1      {Milk, Beer, Diapers}

2      {Bread, Butter, Milk}

3      {Milk, Diapers, Cookies}

4      {Bread, Butter, Cookies}

5      {Beer, Cookies, Diapers}

6      {Milk, Diapers, Bread, Butter}

7      {Bread, Butter, Diapers}

8      {Beer, Diapers}

9      {Milk, Diapers, Bread, Butter}

10     {Beer, Cookies}

What is the maximum number of size 3-itemsets that can be derived from this data set. Note that the itemsets derived may not be in the dataset.

20

35

40

15

None of the above.

### Question 5

2 / 2 pts

Transaction ID Items Bought

1 {Milk, Beer, Diapers}

2 {Bread, Butter, Milk}

3 {Milk, Diapers, Cookies}

4 {Bread, Butter, Cookies}

5 {Beer, Cookies, Diapers}

6 {Milk, Diapers, Bread, Butter}

7 {Bread, Butter, Diapers}

8 {Beer, Diapers}

9 {Milk, Diapers, Bread, Butter}

10 {Beer, Cookies}

What is the confidence of the rule Beer -> Cookies

0.5

### Question 6

2 / 2 pts

Increase in size of a convolutional hidden layer would not necessarily increase the performance of a convolutional network.

True

False

### Question 7

2 / 2 pts

Consider a movie recommendation application, where an user can log in and select movies they like based on their previous preferences.

Which neural network architecture would be suitable to complete this task

Fully-Connected Neural Network.

Convolutional Neural Network.

Recurrent Neural Network.

Restricted Boltzmann Machine.

**Question 8****2 / 2 pts**

What is an autoencoder?

- a neural network that copies its input to its output.
- a neural network that codes itself.
- 
- a neural network that maps an output to an input through a hidden layer.
- a neural network that is trained to attempt to copy its input to its output.

**Question 9****2 / 2 pts**

Which of the following is an application of autoencoders?

- feature learning.
- dimensionality reduction.
- Rule mining
- all of the above.

**Question 10****2 / 2 pts**

Is an autoencoder the same as performing Principal Component Analysis?

- Not at all
- No but conceptually they can be used for the same purpose
- Autoencoders can do dimensionality reduction, but it is non-linear.
- Yes they have the same purpose.

**Question 11****2 / 2 pts**

Review the table below.

Cluster	Tennis	Golf	Baseball	Basketball	Football	Hockey	Total	Entropy	Purity
#1	1	1	0	11	4	676	693		
#2	27	89	333	827	253	33	1562		
#3	326	465	8	105	16	29	949		
Total	354	555	341	943	273	738	3204		

Select the collect entropy set of the confusion matrix.

- Cluster #1: 1.84, Cluster #2: 0.2, Cluster #3: 1.7, Total: 1.44
- Cluster #1: 1.44, Cluster #2: 1.84, Cluster #3: 1.7, Total: 0.2
- Cluster #1: 0.2, Cluster #2: 1.7, Cluster #3: 1.84, Total: 1.44
- Cluster #1: 0.2, Cluster #2: 1.44, Cluster #3: 1.7, Total: 1.84
- None of above

Cluster #1: 0.2, Cluster #2: 1.84, Cluster #3: 1.7, Total: 1.44

**Question 12****2 / 2 pts**

Review the table below.

Cluster	Tennis	Golf	Baseball	Basketball	Football	Hockey	Total	Entropy	Purity
#1	1	1	0	11	4	676	693		
#2	27	89	333	827	253	33	1562		
#3	326	465	8	105	16	29	949		
Total	354	555	341	943	273	738	3204		

Select the collect purity set of the confusion matrix

Cluster #1: 0.98, Cluster #2: 0.53, Cluster #3: 0.49, Total: 0.61

Cluster #1: 0.53, Cluster #2: 0.98, Cluster #3: 0.61, Total: 0.49

Cluster #1: 0.98, Cluster #2: 0.49, Cluster #3: 0.53, Total: 0.61

Cluster #1: 0.53, Cluster #2: 0.49, Cluster #3: 0.61, Total: 0.98

None of above

**Question 13****2 / 2 pts**

Review the following table.

	A	B	C	D	E
A	0	9	3	6	11
B	9	0	7	5	10
C	3	7	0	9	2
D	6	5	9	0	8
E	11	10	2	8	0

Select the correct DBSCAN clustering result with the parameters below

epsilon = 5, Minpts = 2

- cluster1: {A, C, E}, cluster2: {B,D}
- cluster1: {A}, cluster2: {B}, cluster3: {D}, and cluster4: {C, E}
- cluster1: {A, B}, cluster2: {C}, and cluster3: {D}, cluster4: {E}
- cluster1: {A}, cluster2: {B, E}, cluster3: {C}, and cluster4: {D}
- None of above

### Question 14

2 / 2 pts

Select qualities of clusters produced by a good clustering algorithm.

- Intra-cluster distances are minimized
- Inter-cluster distances are maximized
- Number of clusters produced
- A and B are correct

good cluster have strong cohesion within the cluster and maximum distance between them.

### Question 15

2 / 2 pts

What type of clustering is DBSCAN?

Center-based Contiguous Well-separated None of above

DBSCAN is a density-based clustering algorithm

**Question 16****2 / 2 pts**

Review the following table.

Data	X	Y
x1	1	1
x2	6	6
x3	7	7
x4	3	2
x5	6	8
x6	2	1
x7	2	3

Compute the average silhouette score to evaluate K-mean clustering result with the parameters below.

K:2

Initial centroids: centroid1 = (4,3) and centroid2 = (9,7)

Distance measure: Euclidean distance

0.752

**Question 17****2 / 2 pts**

Select all that are not true about DBSCAN.

- it has trouble when the clusters have widely varying densities
- it is very sensitive to noise and to the size of the data
- DBSCAN has a complexity of order  $O(n^3)$
- it can be expensive when the computation of nearest neighbors requires computing all pairwise proximities
- A and B are correct
- B and C are correct
  - DBSCAN is relatively resistant to noise and can handle clusters of arbitrary shapes and sizes. It has complexity order  $O(n \log n)$
- A, C and D are correct

**Question 18****2 / 2 pts**

Select the true statement for Nearest Neighbor classification



Although Nearest Neighbor classification has the data dimensionality issue, it can be solved by the scaling technique



k-NN classifier is a typical lazy learner because it spend so much time building the model even if the 'k' is very small



Drawing a circle is a mandatory task to find neighbor classes regardless of the number of 'K'



Determining the optimal value of 'K' is important for k-NN classifier performance.



None of above

### Question 19

2 / 2 pts

Review the table below.

Instance	X-Axis	Class
i	-3	-
ii	-2	-
iii	-1	+
iv	0	+
v	1	+
vi	2	-

Select all the support vector instances using Non-linear SVM

## Feature Map ( $\{R^1 \rightarrow R^2\}$ )

 i ii iii iv v vi

### Question 20

2 / 2 pts

Review the table below.

Instance	X-Axis	Class
i	-3	-
ii	-2	-
iii	-1	+
iv	0	+
v	1	+
vi	2	-

Select a correct hyper-plane equation of the data below using Non-linear SVM.

## Feature Map ( $\{R^1 \rightarrow R^2\}$ )

- Y=1
- Y=1.5
- Y=2
- Y=2.5
- None of above

Quiz Score: **40** out of 40

# Final Exam

**Due Dec 9 at 1:15am      Points 40      Questions 20**

**Available Dec 6 at 12am - Dec 9 at 2am 3 days**

**Time Limit 75 Minutes**

## Instructions

No collaboration

75 mins

## Attempt History

	<b>Attempt</b>	<b>Time</b>	<b>Score</b>
<b>LATEST</b>	<a href="#"><u>Attempt 1</u></a>	44 minutes	40 out of 40

! Correct answers are hidden.

Score for this quiz: **40** out of 40

Submitted Dec 8 at 6pm

This attempt took 44 minutes.

### Question 1

2 / 2 pts

Transaction ID Items Bought

- 1 {Milk, Beer, Diapers}
- 2 {Bread, Butter, Milk}
- 3 {Milk, Diapers, Cookies}
- 4 {Bread, Butter, Cookies}

- 5        {Beer, Cookies, Diapers}
- 6        {Milk, Diapers, Bread, Butter}
- 7        {Bread, Butter, Diapers}
- 8        {Beer, Diapers}
- 9        {Milk, Diapers, Bread, Butter}
- 10      {Beer, Cookies}

What is the maximum size of frequent itemsets that can be extracted  
(assuming minsup > 0.1)?

---

10

---

9

---

4

---

3

## Question 2

2 / 2 pts

Which of the above statements are true for any A, B, and C?

---

If A → B then B → A.

---

If A → B and B → C then A → C.

---

If A → C then A union B → C.

---

If A union B  $\rightarrow$  C then A  $\rightarrow$  C.

I & II

I, II, & III

I, II, & IV

II & III

none

### Question 3

2 / 2 pts

Example of market basket transactions.

#### Transaction ID Items Bought

1 {a, b, d, e}

2 {b, c, d}

3 {a, b, d, e}

4 {a, c, d, e}

5 {b, c, d, e}

6 {b, d, e}

7 {c, d}

8 {a, b, c}

9      {a, d, e}

10     {b, d}

Compute the confidence for the association rules  $\{a, e\} \rightarrow \{b\}$ .

0.5

#### Question 4

2 / 2 pts

Transaction ID Items Bought

1      {Milk, Beer, Diapers}

2      {Bread, Butter, Milk}

3      {Milk, Diapers, Cookies}

4      {Bread, Butter, Cookies}

5      {Beer, Cookies, Diapers}

6      {Milk, Diapers, Bread, Butter}

7      {Bread, Butter, Diapers}

8      {Beer, Diapers}

9      {Milk, Diapers, Bread, Butter}

10     {Beer, Cookies}

What is the maximum number of size 3-itemsets that can be derived from this data set. Note that the itemsets derived may not be in the dataset.

20

35

40

15

None of the above.

### Question 5

2 / 2 pts

Transaction ID Items Bought

- |   |                                |
|---|--------------------------------|
| 1 | {Milk, Beer, Diapers}          |
| 2 | {Bread, Butter, Milk}          |
| 3 | {Milk, Diapers, Cookies}       |
| 4 | {Bread, Butter, Cookies}       |
| 5 | {Beer, Cookies, Diapers}       |
| 6 | {Milk, Diapers, Bread, Butter} |
| 7 | {Bread, Butter, Diapers}       |
| 8 | {Beer, Diapers}                |

9 {Milk, Diapers, Bread, Butter}

10 {Beer, Cookies}

What is the confidence of the rule Beer -> Cookies

0.5

### Question 6

2 / 2 pts

Increase in size of a convolutional hidden layer would not necessarily increase the performance of a convolutional network.

True

False

### Question 7

2 / 2 pts

Consider a movie recommendation application, where an user can log in and select movies they like based on their previous preferences.

Which neural network architecture would be suitable to complete this task

Fully-Connected Neural Network.

Convolutional Neural Network.

Recurrent Neural Network.

Restricted Boltzmann Machine.

**Question 8****2 / 2 pts**

What is an autoencoder?

- a neural network that copies its input to its output.
- a neural network that codes itself.
- 
- a neural network that maps an output to an input through a hidden layer.
- a neural network that is trained to attempt to copy its input to its output.

**Question 9****2 / 2 pts**

Which of the following is an application of autoencoders?

- feature learning.
- dimensionality reduction.
- Rule mining
- all of the above.

**Question 10****2 / 2 pts**

Is an autoencoder the same as performing Principal Component Analysis?

- Not at all
- No but conceptually they can be used for the same purpose
- Autoencoders can do dimensionality reduction, but it is non-linear.
- Yes they have the same purpose.

**Question 11****2 / 2 pts**

Review the table below.

Cluster	Tennis	Golf	Baseball	Basketball	Football	Hockey	Total	Entropy	Purity
#1	1	1	0	11	4	676	693		
#2	27	89	333	827	253	33	1562		
#3	326	465	8	105	16	29	949		
Total	354	555	341	943	273	738	3204		

Select the collect entropy set of the confusion matrix.

- Cluster #1: 1.84, Cluster #2: 0.2, Cluster #3: 1.7, Total: 1.44
- Cluster #1: 1.44, Cluster #2: 1.84, Cluster #3: 1.7, Total: 0.2
- Cluster #1: 0.2, Cluster #2: 1.7, Cluster #3: 1.84, Total: 1.44
- Cluster #1: 0.2, Cluster #2: 1.44, Cluster #3: 1.7, Total: 1.84
- None of above

Cluster #1: 0.2, Cluster #2: 1.84, Cluster #3: 1.7, Total: 1.44

**Question 12****2 / 2 pts**

Review the table below.

Cluster	Tennis	Golf	Baseball	Basketball	Football	Hockey	Total	Entropy	Purity
#1	1	1	0	11	4	676	693		
#2	27	89	333	827	253	33	1562		
#3	326	465	8	105	16	29	949		
Total	354	555	341	943	273	738	3204		

Select the collect purity set of the confusion matrix

Cluster #1: 0.98, Cluster #2: 0.53, Cluster #3: 0.49, Total: 0.61

Cluster #1: 0.53, Cluster #2: 0.98, Cluster #3: 0.61, Total: 0.49

Cluster #1: 0.98, Cluster #2: 0.49, Cluster #3: 0.53, Total: 0.61

Cluster #1: 0.53, Cluster #2: 0.49, Cluster #3: 0.61, Total: 0.98

None of above

**Question 13****2 / 2 pts**

Review the following table.

	A	B	C	D	E
A	0	9	3	6	11
B	9	0	7	5	10
C	3	7	0	9	2
D	6	5	9	0	8
E	11	10	2	8	0

Select the correct DBSCAN clustering result with the parameters below

epsilon = 5, Minpts = 2

- cluster1: {A, C, E}, cluster2: {B,D}
- cluster1: {A}, cluster2: {B}, cluster3: {D}, and cluster4: {C, E}
- cluster1: {A, B}, cluster2: {C}, and cluster3: {D}, cluster4: {E}
- cluster1: {A}, cluster2: {B, E}, cluster3: {C}, and cluster4: {D}
- None of above

### Question 14

2 / 2 pts

Select qualities of clusters produced by a good clustering algorithm.

- Intra-cluster distances are minimized
- Inter-cluster distances are maximized
- Number of clusters produced
- A and B are correct

good cluster have strong cohesion within the cluster and maximum distance between them.

### Question 15

2 / 2 pts

What type of clustering is DBSCAN?

Center-based Contiguous Well-separated None of above

DBSCAN is a density-based clustering algorithm

**Question 16****2 / 2 pts**

Review the following table.

Data	X	Y
x1	1	1
x2	6	6
x3	7	7
x4	3	2
x5	6	8
x6	2	1
x7	2	3

Compute the average silhouette score to evaluate K-mean clustering result with the parameters below.

K:2

Initial centroids: centroid1 = (4,3) and centroid2 = (9,7)

Distance measure: Euclidean distance

0.752

**Question 17****2 / 2 pts**

Select all that are not true about DBSCAN.

- it has trouble when the clusters have widely varying densities
- it is very sensitive to noise and to the size of the data
- DBSCAN has a complexity of order  $O(n^3)$
- it can be expensive when the computation of nearest neighbors requires computing all pairwise proximities
- A and B are correct
- B and C are correct
  - DBSCAN is relatively resistant to noise and can handle clusters of arbitrary shapes and sizes. It has complexity order  $O(n \log n)$
- A, C and D are correct

**Question 18****2 / 2 pts**

Select the true statement for Nearest Neighbor classification



Although Nearest Neighbor classification has the data dimensionality issue, it can be solved by the scaling technique



k-NN classifier is a typical lazy learner because it spend so much time building the model even if the 'k' is very small



Drawing a circle is a mandatory task to find neighbor classes regardless of the number of 'K'



Determining the optimal value of 'K' is important for k-NN classifier performance.



None of above

### Question 19

2 / 2 pts

Review the table below.

Instance	X-Axis	Class
i	-3	-
ii	-2	-
iii	-1	+
iv	0	+
v	1	+
vi	2	-

Select all the support vector instances using Non-linear SVM

## Feature Map ( $\{R^1 \rightarrow R^2\}$ )

 i ii iii iv v vi

### Question 20

2 / 2 pts

Review the table below.

Instance	X-Axis	Class
i	-3	-
ii	-2	-
iii	-1	+
iv	0	+
v	1	+
vi	2	-

Select a correct hyper-plane equation of the data below using Non-linear SVM.

## Feature Map ( $\{R^1 \rightarrow R^2\}$ )

- Y=1
- Y=1.5
- Y=2
- Y=2.5
- None of above

Quiz Score: **40** out of 40

# Final Exam

**Due Dec 9 at 1:15am      Points 40      Questions 20**

**Available Dec 6 at 12am - Dec 9 at 2am 3 days**

**Time Limit 75 Minutes**

## Instructions

No collaboration

75 mins

## Attempt History

	<b>Attempt</b>	<b>Time</b>	<b>Score</b>
<b>LATEST</b>	<a href="#"><u>Attempt 1</u></a>	54 minutes	40 out of 40

! Correct answers are hidden.

Score for this quiz: **40** out of 40

Submitted Dec 8 at 5:57pm

This attempt took 54 minutes.

### Question 1

2 / 2 pts

Apriori pruning based on support is a greedy strategy but is not optimal

True

False

### Question 2

2 / 2 pts

Example of market basket transactions.

Transaction ID Items Bought

- |    |              |
|----|--------------|
| 1  | {a, b, d, e} |
| 2  | {b, c, d}    |
| 3  | {a, b, d, e} |
| 4  | {a, c, d, e} |
| 5  | {b, c, d, e} |
| 6  | {b, d, e}    |
| 7  | {c, d}       |
| 8  | {a, b, c}    |
| 9  | {a, d, e}    |
| 10 | {b, d}       |

Compute the support for itemset {b,c,d}.

0.2

**Question 3**

**2 / 2 pts**

Example of market basket transactions.

**Transaction ID Items Bought**

1 {a, b, d, e}

2 {b, c, d}

3 {a, b, d, e}

4 {a, c, d, e}

5 {b, c, d, e}

6 {b, d, e}

7 {c, d}

8 {a, b, c}

9 {a, d, e}

10 {b, d}

If the minimum support is set at 40%, how many frequent 3-itemsets will be found? Note that if a 3 itemset is a subset of a larger itemset, it counts as one occurrence.

2

**Question 4**

2 / 2 pts

Select the reason why Apriori pruning in the search for frequent itemsets works.

- support count is monotonic with respect to itemsets.
- we search in transaction ID order.
- support count diverges as we add to the itemset.
- support count is anti-monotonic with respect to itemsets.
- it is an excellent heuristic, but it does not work 100% of the time.

**Question 5****2 / 2 pts****Transaction ID Items Bought**

- |    |                                |
|----|--------------------------------|
| 1  | {Milk, Beer, Diapers}          |
| 2  | {Bread, Butter, Milk}          |
| 3  | {Milk, Diapers, Cookies}       |
| 4  | {Bread, Butter, Cookies}       |
| 5  | {Beer, Cookies, Diapers}       |
| 6  | {Milk, Diapers, Bread, Butter} |
| 7  | {Bread, Butter, Diapers}       |
| 8  | {Beer, Diapers}                |
| 9  | {Milk, Diapers, Bread, Butter} |
| 10 | {Beer, Cookies}                |

Find all itemsets (of size 3 or larger) that has the largest support.

{Milk, Beer, Diapers}

{Bread, Butter, Milk}

{Bread, Butter, Diapers}

{Milk, Diapers, Bread}

None of the above.

All of the above.

## Question 6

2 / 2 pts

Usage of deep learning is predicated with availability of large amounts of data. Is the data volume the only determining factor to check whether we can use deep learning ?

Yes, lower volumes lead to overfitting

No, balance among classes in the input data is the only determining factor  
volume does not matter

No, balance among classes in the input data is a co-factor along with the  
data volume

None of the above

**Question 7****2 / 2 pts**

Consider a movie recommendation application, where an user can log in and select movies they like based on their previous preferences.

Which neural network architecture would be suitable to complete this task

- 
- Fully-Connected Neural Network.
  - Convolutional Neural Network.
  - Recurrent Neural Network.
  - Restricted Boltzmann Machine.
- 

**Question 8****2 / 2 pts**

Which of the following is an application of autoencoders?

- 
- feature learning.
  - dimensionality reduction.
  - Rule mining
  - all of the above.
- 

**Question 9****2 / 2 pts**

What is an autoencoder?

- a neural network that copies its input to its output.
- a neural network that codes itself.
- 
- a neural network that maps an output to an input through a hidden layer.
- a neural network that is trained to attempt to copy its input to its output.

**Question 10****2 / 2 pts**

Increase in size of a convolutional hidden layer would not necessarily increase the performance of a convolutional network.

- True
- False

**Question 11****2 / 2 pts**

Review the table below.

Cluster	Tennis	Golf	Baseball	Basketball	Football	Hockey	Total	Entropy	Purity
#1	1	1	0	11	4	676	693		
#2	27	89	333	827	253	33	1562		
#3	326	465	8	105	16	29	949		
Total	354	555	341	943	273	738	3204		

Select the collect entropy set of the confusion matrix.

- Cluster #1: 1.84, Cluster #2: 0.2, Cluster #3: 1.7, Total: 1.44

Cluster #1: 1.44, Cluster #2: 1.84, Cluster #3: 1.7, Total: 0.2

Cluster #1: 0.2, Cluster #2: 1.7, Cluster #3: 1.84, Total: 1.44

Cluster #1: 0.2, Cluster #2: 1.44, Cluster #3: 1.7, Total: 1.84

None of above

Cluster #1: 0.2, Cluster #2: 1.84, Cluster #3: 1.7, Total: 1.44

## Question 12

2 / 2 pts

Select qualities of clusters produced by a good clustering algorithm.

Intra-cluster distances are minimized

Inter-cluster distances are maximized

Number of clusters produced

A and B are correct

good cluster have strong cohesion within the cluster and maximum distance between them.

## Question 13

2 / 2 pts

Select all statements that are true for Hierarchical clustering.

Although it requires high storage, high computing is not required

Robust to noise Robust to outliers

It is not easy to decide on the number of cluster because the complicated dendrogram disturbs this decision

 None of above

### Question 14

2 / 2 pts

Review the following table.

Data	X	Y
A	1	2
B	2	2
C	2	3
D	8	7
E	8	8
F	8	2

Select the correct number of core points and DBSCAN clustering result with the parameters below.

Eps: 5

Minimum samples (MinPts): 3

Distance measurement: Euclidean distance

 5, cluster1: {A,B,C}, cluster2: {D,E,F} 6, cluster1: {A,B}, cluster2: {C,D,E}, cluster3: {F} 4, cluster1: {A,B,C}, cluster2: {D, E, F}

- 5, cluster1: {A}, cluster2: {B}, cluster3: {C}, cluster4: {D}, cluster5: {E}
- None of above

**Question 15****2 / 2 pts**

Review the following table.

Data	X	Y
x1	1	1
x2	6	6
x3	7	7
x4	3	2
x5	6	8
x6	2	1
x7	2	3

Compute the sum of squared error (SSE) to evaluate K-mean clustering result with the parameters below.

K:2

Initial centroids: centroid1 = (4,3) and centroid2 = (9,7)

Distance measure: Euclidean distance

7.4167

**Question 16****2 / 2 pts**

Select all that are not true about DBSCAN.

- it has trouble when the clusters have widely varying densities
- it is very sensitive to noise and to the size of the data
- DBSCAN has a complexity of order  $O(n^3)$
- it can be expensive when the computation of nearest neighbors requires computing all pairwise proximities
- A and B are correct
- B and C are correct
  - DBSCAN is relatively resistant to noise and can handle clusters of arbitrary shapes and sizes. It has complexity order  $O(n \log n)$
- A, C and D are correct

**Question 17****2 / 2 pts**

Select all different types of clusters.

- Well-separated
- Center-based
- Contiguous
- Density-based
- All of the above

all are type of clustering

**Question 18**

2 / 2 pts

Review the table below.

Instance	X-Axis	Class
i	-3	-
ii	-2	-
iii	-1	+
iv	0	+
v	1	+
vi	2	-

Select all the support vector instances using Non-linear SVM

Feature Map ( $\{R^1 \rightarrow R^2\}$ )

---

 i

---

 ii

---

 iii

---

 iv

---

 v

---

 vi

**Question 19****2 / 2 pts**

Review the table below.

A	B	C	Class
F	F	F	Y
F	F	T	N
F	T	T	N
F	T	T	N
F	F	T	Y
T	F	T	Y
T	F	T	N
T	F	T	N
T	T	T	Y
T	F	T	Y

Select the correct prediction class for a test sample (A=F, B=T, C=F) using the naïve Bayes approach.

---

 'Y'

---

 'N'

---

 Anyone**Question 20****2 / 2 pts**

Review the table below.

Rainy	Sunshine		Wind		Atmosphere		Temperature	
	Yes	No	Yes	No	High	Low	High	Low
Yes	2	6	6	2	8	0	5	3
No	8	4	2	10	2	10	6	6

Is it raining if the weather condition is *sunshine, no wind, high atmosphere*, and *low temperature* (*Hint: you should use Naïve Bayes*)?

Yes

No

Quiz Score: **40** out of 40

# Final Exam

**Due** Dec 8 at 1:15am      **Points** 40      **Questions** 20

**Available** Dec 6 at 12am - Dec 8 at 2am 2 days

**Time Limit** 75 Minutes

## Instructions

No collaboration

75 mins

## Attempt History

	<b>Attempt</b>	<b>Time</b>	<b>Score</b>
<b>LATEST</b>	<a href="#"><u>Attempt 1</u></a>	50 minutes	38 out of 40

! Correct answers are hidden.

Score for this quiz: **38** out of 40

Submitted Dec 6 at 2:37pm

This attempt took 50 minutes.

<b>Question 1</b>		<b>2 / 2 pts</b>
Example of market basket transactions.		
Transaction ID Items Bought		
1	{a, b, d, e}	
2	{b, c, d}	
3	{a, b, d, e}	

- 4      {a, c, d, e}
- 5      {b, c, d, e}
- 6      {b, d, e}
- 7      {c, d}
- 8      {a, b, c}
- 9      {a, d, e}
- 10     {b, d}

Compute the support for itemset {b,c,d}.

0.2

## Question 2

2 / 2 pts

Example of market basket transactions.

### Transaction ID Items Bought

- 1      {a, b, d, e}
- 2      {b, c, d}
- 3      {a, b, d, e}
- 4      {a, c, d, e}

5      {b, c, d, e}

6      {b, d, e}

7      {c, d}

8      {a, b, c}

9      {a, d, e}

10     {b, d}

Compute the confidence for the association rules  $\{a, e\} \rightarrow \{b\}$ .

0.5

### Question 3

2 / 2 pts

Transaction ID Items Bought

1      {Milk, Beer, Diapers}

2      {Bread, Butter, Milk}

3      {Milk, Diapers, Cookies}

4      {Bread, Butter, Cookies}

5      {Beer, Cookies, Diapers}

6      {Milk, Diapers, Bread, Butter}

- 7 {Bread, Butter, Diapers}
- 8 {Beer, Diapers}
- 9 {Milk, Diapers, Bread, Butter}
- 10 {Beer, Cookies}

What is the confidence of the rule Beer  $\rightarrow$  Cookies

0.5

#### Question 4

2 / 2 pts

Example of market basket transactions.

#### Transaction ID Items Bought

- 1 {a, b, d, e}
- 2 {b, c, d}
- 3 {a, b, d, e}
- 4 {a, c, d, e}
- 5 {b, c, d, e}
- 6 {b, d, e}
- 7 {c, d}

8            {a, b, c}

9            {a, d, e}

10          {b, d}

If the minimum support is set at 40%, how many frequent 3-itemsets will be found? Note that if a 3 itemset is a subset of a larger itemset, it counts as one occurrence.

2

### Question 5

2 / 2 pts

Transaction ID Items Bought

1            {Milk, Beer, Diapers}

2            {Bread, Butter, Milk}

3            {Milk, Diapers, Cookies}

4            {Bread, Butter, Cookies}

5            {Beer, Cookies, Diapers}

6            {Milk, Diapers, Bread, Butter}

7            {Bread, Butter, Diapers}

8            {Beer, Diapers}

9 {Milk, Diapers, Bread, Butter}

10 {Beer, Cookies}

What is the maximum number of size 3-itemsets that can be derived from this data set. Note that the itemsets derived may not be in the dataset.

20

35

40

15

None of the above.

### Question 6

2 / 2 pts

Which of the following is an application of autoencoders?

feature learning.

dimensionality reduction.

Rule mining

all of the above.

### Question 7

2 / 2 pts

The number of neurons in the output layer should match the number of classes (Where the number of classes is greater than 2) in a supervised learning task.

---

True

---

False

**Question 8****2 / 2 pts**

What is an autoencoder?

---

a neural network that copies its input to its output.

---

a neural network that codes itself.

---

a neural network that maps an output to an input through a hidden layer.

---

a neural network that is trained to attempt to copy its input to its output.

**Question 9****2 / 2 pts**

Consider a movie recommendation application, where an user can log in and select movies they like based on their previous preferences.

Which neural network architecture would be suitable to complete this task

---

Fully-Connected Neural Network.

---

Convolutional Neural Network.

---

- Recurrent Neural Network.
- 
- Restricted Boltzmann Machine.

Incorrect

**Question 10****0 / 2 pts**

Increase in size of a convolutional hidden layer would not necessarily increase the performance of a convolutional network.

- 
- True
- 
- False

**Question 11****2 / 2 pts**

Select different aspects of cluster validation.

- 
- Determining the clustering tendency of a set of data
- 
- To improve the complexity of the algorithm
- 
- Determining the 'correct' number of clusters
- 
- None of the above

correct these are some of the aspects of cluster validation!

**Question 12****2 / 2 pts**

What is the purpose of cluster analysis?

- To avoid finding patterns in noise
- To compare clustering algorithms
- To compare two sets of clusters
- all of the above
- None of the above

we do cluster validity to avoid clustering noise and find a suitable algorithm for our data

**Question 13****2 / 2 pts**

Select qualities of clusters produced by a good clustering algorithm.

- Intra-cluster distances are minimized
- Inter-cluster distances are maximized
- Number of clusters produced
- A and B are correct

good cluster have strong cohesion within the cluster and maximum distance between them.

**Question 14****2 / 2 pts**

Select all statements that are true for Hierarchical clustering.

Although it requires high storage, high computing is not required

Robust to noise

Robust to outliers

It is not easy to decide on the number of cluster because the complicated dendrogram disturbs this decision

None of above

**Question 15****2 / 2 pts**

Review the following table.

Data	X	Y
A	1	2
B	2	2
C	2	3
D	8	7
E	8	8
F	8	2

Select the correct number of core points and DBSCAN clustering result with the parameters below.

Eps: 5

Minimum samples (MinPts):3

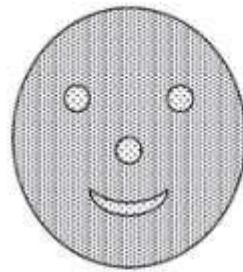
### Distance measurement: Euclidean distance

- 5, cluster1: {A,B,C}, cluster2: {D,E,F}
- 6, cluster1: {A,B}, cluster2: {C,D,E}, cluster3: {F}
- 4, cluster1: {A,B,C}, cluster2: {D, E, F}
- 5, cluster1: {A}, cluster2: {B}, cluster3: {C}, cluster4: {D}, cluster5: {E}
- None of above

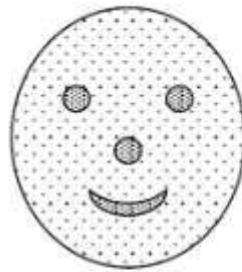
### Question 16

2 / 2 pts

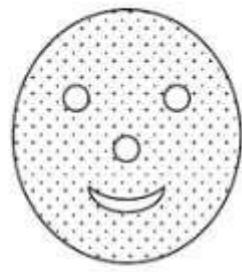
Review the image below.



(a)



(b)



(c)



(d)

**Darkness or number of dots represent density. Lines are used only to distinguish regions and do not represent points.**

If you want to find the patterns represented by the nose, eyes, and mouth using k-Means clustering, select all figures that be well-clustered?

 (a) (b) (c)

(d) None of above**Question 17****2 / 2 pts**

Select all that are not true about DBSCAN.

- it has trouble when the clusters have widely varying densities
- it is very sensitive to noise and to the size of the data
- DBSCAN has a complexity of order  $O(n^3)$
- 
- it can be expensive when the computation of nearest neighbors requires computing all pairwise proximities
- A and B are correct
- B and C are correct

DBSCAN is relatively resistant to noise and can handle clusters of arbitrary shapes and sizes. It has complexity order  $O(n \log n)$

- A, C and D are correct

**Question 18****2 / 2 pts**

Review the table below.

Instance	X-Axis	Class
i	-3	-
ii	-2	-
iii	-1	+
iv	0	+
v	1	+
vi	2	-

Select a correct hyper-plane equation of the data below using Non-linear SVM.

## Feature Map ( $\{R^1 \rightarrow R^2\}$ )

Y=1

Y=1.5

Y=2

Y=2.5

None of above

### Question 19

2 / 2 pts

Review the table below.

Instance	X-Axis	Class
i	-3	-
ii	-2	-
iii	-1	+
iv	0	+
v	1	+
vi	2	-

Select all the support vector instances using Non-linear SVM

## Feature Map ( $\{R^1 \rightarrow R^2\}$ )

i

ii

iii

iv

v

vi

### Question 20

2 / 2 pts

Review the table below.

A	B	C	Class
F	F	F	Y
F	F	T	N
F	T	T	N
F	T	T	N
F	F	T	Y
T	F	T	Y
T	F	T	N
T	F	T	N
T	T	T	Y
T	F	T	Y

Select the correct prediction class for a test sample (A=F, B=T, C=F) using the naïve Bayes approach.

'Y'

'N'

Anyone

Quiz Score: **38** out of 40

# Final Exam

**Due** Dec 9 at 1:15am      **Points** 40      **Questions** 20

**Available** Dec 6 at 12am - Dec 9 at 2am 3 days

**Time Limit** 75 Minutes

## Instructions

No collaboration

75 mins

## Attempt History

	Attempt	Time	Score
LATEST	<a href="#"><u>Attempt 1</u></a>	39 minutes	40 out of 40

❗ Correct answers are hidden.

Score for this quiz: **40** out of 40

Submitted Dec 8 at 5:10pm

This attempt took 39 minutes.

Question 1	2 / 2 pts
Select the reason why Apriori pruning in the search for frequent itemsets works.	
<input type="radio"/> support count is monotonic with respect to itemsets.	
<input type="radio"/> we search in transaction ID order.	
<input type="radio"/> support count diverges as we add to the itemset.	
<input checked="" type="radio"/> support count is anti-monotonic with respect to itemsets.	

- it is an excellent heuristic, but it does not work 100% of the time.

**Question 2****2 / 2 pts**

Example of market basket transactions.

Transaction ID Items Bought

1 {a, b, d, e}

2 {b, c, d}

3 {a, b, d, e}

4 {a, c, d, e}

5 {b, c, d, e}

6 {b, d, e}

7 {c, d}

8 {a, b, c}

9 {a, d, e}

10 {b, d}

Compute the support for itemset {b,c,d}.

0.2

**Question 3**

**2 / 2 pts**

## Transaction ID Items Bought

- 1 {Milk, Beer, Diapers}
- 2 {Bread, Butter, Milk}
- 3 {Milk, Diapers, Cookies}
- 4 {Bread, Butter, Cookies}
- 5 {Beer, Cookies, Diapers}
- 6 {Milk, Diapers, Bread, Butter}
- 7 {Bread, Butter, Diapers}
- 8 {Beer, Diapers}
- 9 {Milk, Diapers, Bread, Butter}
- 10 {Beer, Cookies}

What is the confidence of the rule Beer  $\rightarrow$  Cookies

0.5

**Question 4****2 / 2 pts**

Example of market basket transactions.

## Transaction ID Items Bought

1 {a, b, d, e}

2 {b, c, d}

3 {a, b, d, e}

4 {a, c, d, e}

5 {b, c, d, e}

6 {b, d, e}

7 {c, d}

8 {a, b, c}

9 {a, d, e}

10 {b, d}

If the minimum support is set at 40%, how many frequent 3-itemsets will be found? Note that if a 3 itemset is a subset of a larger itemset, it counts as one occurrence.

2

**Question 5****2 / 2 pts**

Apriori pruning based on support is a greedy strategy but is not optimal

True False**Question 6****2 / 2 pts**

Usage of deep learning is predicated with availability of large amounts of data. Is the data volume the only determining factor to check whether we can use deep learning ?

 Yes, lower volumes lead to overfitting

No, balance among classes in the input data is the only determining factor volume does not matter

No, balance among classes in the input data is a co-factor along with the data volume

 None of the above**Question 7****2 / 2 pts**

The number of neurons in the output layer should match the number of classes (Where the number of classes is greater than 2) in a supervised learning task.

 True False

**Question 8****2 / 2 pts**

Is an autoencoder the same as performing Principal Component Analysis?

- 
- Not at all
- 
- No but conceptually they can be used for the same purpose
- 
- Autoencoders can do dimensionality reduction, but it is non-linear.
- 
- Yes they have the same purpose.

**Question 9****2 / 2 pts**

Consider a movie recommendation application, where an user can log in and select movies they like based on their previous preferences.

Which neural network architecture would be suitable to complete this task

- 
- Fully-Connected Neural Network.
- 
- Convolutional Neural Network.
- 
- Recurrent Neural Network.
- 
- Restricted Boltzmann Machine.

**Question 10****2 / 2 pts**

Increase in size of a convolutional hidden layer would not necessarily increase the performance of a convolutional network.

True

False

### Question 11

2 / 2 pts

Review the table below.

Cluster	Tennis	Golf	Baseball	Basketball	Football	Hockey	Total	Entropy	Purity
#1	1	1	0	11	4	676	693		
#2	27	89	333	827	253	33	1562		
#3	326	465	8	105	16	29	949		
Total	354	555	341	943	273	738	3204		

Select the collect purity set of the confusion matrix

Cluster #1: 0.98, Cluster #2: 0.53, Cluster #3: 0.49, Total: 0.61

Cluster #1: 0.53, Cluster #2: 0.98, Cluster #3: 0.61, Total: 0.49

Cluster #1: 0.98, Cluster #2: 0.49, Cluster #3: 0.53, Total: 0.61

Cluster #1: 0.53, Cluster #2: 0.49, Cluster #3: 0.61, Total: 0.98

None of above

### Question 12

2 / 2 pts

Review the following table.

Data	X	Y
x1	1	1
x2	6	6
x3	7	7
x4	3	2
x5	6	8
x6	2	1
x7	2	3

Select the correct pair of final centroid and clustering set using K-mean clustering with the parameters below.

K:2

Initial centroids: centroid1 = (4,3) and centroid2 = (9,7)

Distance measure: Euclidean distance



Cluster 1: (x1, x4, x6, x7) with centroid=(2, 1.75), Cluster 2: (x2, x3, x5) with centroid=(6.33, 7)



Cluster 1: (x1, x4, x6, x7) with centroid=(1.75, 2), Cluster 2: (x2, x3, x5) with centroid=(6.33, 7)



Cluster 1: (x1, x4, x6, x7) with centroid=(2, 1.75), Cluster 2: (x2, x3, x5) with centroid=(7, 6.33)



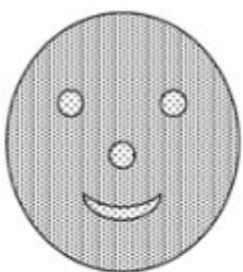
Cluster 1: (x1, x2, x4, x6, x7) with centroid=(2, 1.75), Cluster 2: (x3, x5) with centroid=(6.33, 7)



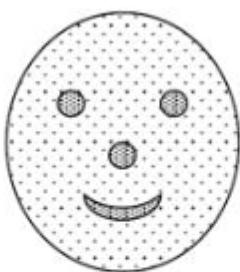
Cluster 1: (x1, x2, x4, x6, x7) with centroid=(2, 1.75), Cluster 2: (x3, x5) with centroid=(7, 6.33)

**Question 13****2 / 2 pts**

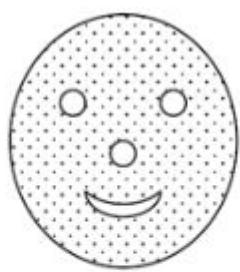
Review the image below.



(a)



(b)



(c)



(d)

**Darkness or number of dots represent density. Lines are used only to distinguish regions and do not represent points.**

If you want to find the patterns represented by the nose, eyes, and mouth using k-Means clustering, select all figures that be well-clustered?

 (a) (b) (c) (d) None of above**Question 14****2 / 2 pts**

Select qualities of clusters produced by a good clustering algorithm.

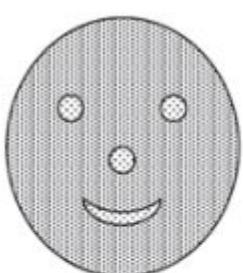
- Intra-cluster distances are minimized
- Inter-cluster distances are maximized
- Number of clusters produced
- A and B are correct

good cluster have strong cohesion within the cluster and maximum distance between them.

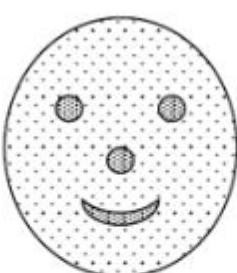
### Question 15

2 / 2 pts

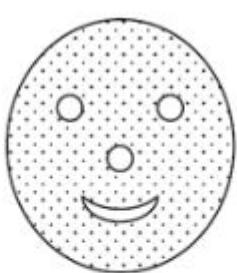
Review the image below.



(a)



(b)



(c)



(d)

**Darkness or number of dots represent density. Lines are used only to distinguish regions and do not represent points.**

If we want to find the patterns represented by the nose, eyes, and mouth using single linkage hierarchical clustering, select all figures that be well-clustered?

(a)

(b)

(c)

(d) None of above**Question 16****2 / 2 pts**

Select all different type of clusters.

 Well-separated Center-based Contiguous Density-based All of the above

all are type of clustering

**Question 17****2 / 2 pts**

Review the following table.

	A	B	C	D	E
A	0	9	3	6	11
B	9	0	7	5	10
C	3	7	0	9	2
D	6	5	9	0	8
E	11	10	2	8	0

Select the correct DBSCAN clustering result with the parameters below

epsilon = 5, Minpts = 2

- cluster1: {A, C, E}, cluster2: {B,D}
- cluster1: {A}, cluster2: {B}, cluster3: {D}, and cluster4: {C, E}
- cluster1: {A, B}, cluster2: {C}, and cluster3: {D}, cluster4: {E}
- cluster1: {A}, cluster2: {B, E}, cluster3: {C}, and cluster4: {D}
- None of above

### Question 18

2 / 2 pts

Review the table below.

A	B	C	Class
F	F	F	Y
F	F	T	N
F	T	T	N
F	T	T	N
F	F	T	Y
T	F	T	Y
T	F	T	N
T	F	T	N
T	T	T	Y
T	F	T	Y

Select the correct prediction class for a test sample (A=F, B=T, C=F) using the naïve Bayes approach.

'Y'

'N'

Anyone

### Question 19

2 / 2 pts

Select the true statement for Nearest Neighbor classification

Although Nearest Neighbor classification has the data dimensionality issue, it can be solved by the scaling technique

k-NN classifier is a typical lazy learner because it spend so much time building the model even if the 'k' is very small

Drawing a circle is a mandatory task to find neighbor classes regardless of the number of 'K'

Determining the optimal value of 'K' is important for k-NN classifier performance.

None of above

### Question 20

2 / 2 pts

Review the table below.

Instance	X-Axis	Class
i	-3	-
ii	-2	-
iii	-1	+
iv	0	+
v	1	+
vi	2	-

Select all the support vector instances using Non-linear SVM

Feature Map ( $\{R^1 \rightarrow R^2\}$ )

 i

<input checked="" type="checkbox"/> ii
<input checked="" type="checkbox"/> iii
<input type="checkbox"/> iv
<input checked="" type="checkbox"/> v
<input checked="" type="checkbox"/> vi

Quiz Score: **40** out of 40

# Final Exam

**Due Dec 9 at 1:15am      Points 40      Questions 20**

**Available** Dec 6 at 12am - Dec 9 at 2am 3 days

**Time Limit** 75 Minutes

## Instructions

No collaboration

75 mins

## Attempt History

	<b>Attempt</b>	<b>Time</b>	<b>Score</b>
<b>LATEST</b>	<a href="#"><u>Attempt 1</u></a>	41 minutes	40 out of 40

! Correct answers are hidden.

Score for this quiz: **40** out of 40

Submitted Dec 8 at 6:17pm

This attempt took 41 minutes.

<b>Question 1</b>	<b>2 / 2 pts</b>
Apriori pruning based on support is a greedy strategy but is not optimal	
<input type="radio"/> True	
<input checked="" type="radio"/> False	

**Question 2****2 / 2 pts**

Which of the above statements are true for any A, B, and C?

- If  $A \rightarrow B$  then  $B \rightarrow A$ .
- If  $A \rightarrow B$  and  $B \rightarrow C$  then  $A \rightarrow C$ .
- If  $A \rightarrow C$  then  $A \cup B \rightarrow C$ .
- If  $A \cup B \rightarrow C$  then  $A \rightarrow C$ .
- I & II
- I, II, & III
- I, II, & IV
- II & III
- none

**Question 3****2 / 2 pts**

Example of market basket transactions.

Transaction ID Items Bought

- |    |              |
|----|--------------|
| 1  | {a, b, d, e} |
| 2  | {b, c, d}    |
| 3  | {a, b, d, e} |
| 4  | {a, c, d, e} |
| 5  | {b, c, d, e} |
| 6  | {b, d, e}    |
| 7  | {c, d}       |
| 8  | {a, b, c}    |
| 9  | {a, d, e}    |
| 10 | {b, d}       |

Compute the confidence for the association rules {a, e} -> {b}.

0.5

#### Question 4

2 / 2 pts

Example of market basket transactions.

**Transaction ID Items Bought**

1 {a, b, d, e}

2 {b, c, d}

3 {a, b, d, e}

4 {a, c, d, e}

5 {b, c, d, e}

6 {b, d, e}

7 {c, d}

8 {a, b, c}

9 {a, d, e}

10 {b, d}

If the minimum support is set at 40%, how many frequent 3-itemsets will be found? Note that if a 3 itemset is a subset of a larger itemset, it counts as one occurrence.

2

**Question 5****2 / 2 pts**

### Example of market basket transactions.

#### Transaction ID Items Bought

1 {a, b, d, e}

2 {b, c, d}

3 {a, b, d, e}

4 {a, c, d, e}

5 {b, c, d, e}

6 {b, d, e}

7 {c, d}

8 {a, b, c}

9 {a, d, e}

10 {b, d}

Compute the support for itemset {b,c,d}.

0.2

**Question 6****2 / 2 pts**

Which of the following is an application of autoencoders?

- 
- feature learning.
- 
- dimensionality reduction.
- 
- Rule mining
- 
- all of the above.

**Question 7****2 / 2 pts**

What is an autoencoder?

- 
- a neural network that copies its input to its output.
- 
- a neural network that codes itself.
- 
- 
- a neural network that maps an output to an input through a hidden layer.
- 
- 
- a neural network that is trained to attempt to copy its input to its output.

**Question 8****2 / 2 pts**

The number of neurons in the output layer should match the number of classes (Where the number of classes is greater than 2) in a supervised learning task.

 True False**Question 9****2 / 2 pts**

Select all that is true about Restricted Boltzman Machine (RBM)

 Can be used for recommendation systems RBM training is often probabilistic RBM is a kind of Deep Belief Network None of the above.**Question 10****2 / 2 pts**

Consider a movie recommendation application, where an user can log in and select movies they like based on their previous preferences.

Which neural network architecture would be suitable to complete this task

- Fully-Connected Neural Network.
- Convolutional Neural Network.
- Recurrent Neural Network.
- Restricted Boltzmann Machine.

### Question 11

2 / 2 pts

Review the following table.

Data	X	Y
x1	1	1
x2	6	6
x3	7	7
x4	3	2
x5	6	8
x6	2	1
x7	2	3

Compute the average silhouette score to evaluate K-mean clustering result with the parameters below.

K:2

Initial centroids: centroid1 = (4,3) and centroid2 = (9,7)

Distance measure: Euclidean distance

0.752

### Question 12

2 / 2 pts

Review the following table.

Data	X	Y
A	1	2
B	2	2
C	2	3
D	8	7
E	8	8
F	8	2

Select the correct number of core points and DBSCAN clustering result with the parameters below.

Eps: 5

Minimum samples (MinPts):3

Distance measurement: Euclidean distance

- 5, cluster1: {A,B,C}, cluster2: {D,E,F}
- 6, cluster1: {A,B}, cluster2: {C,D,E}, cluster3: {F}
- 4, cluster1: {A,B,C}, cluster2: {D, E, F}



5, cluster1: {A}, cluster2: {B}, cluster3: {C}, cluster4: {D}, cluster5: {E}



None of above

### Question 13

2 / 2 pts

Review the following table.

Data	X	Y
x1	1	1
x2	6	6
x3	7	7
x4	3	2
x5	6	8
x6	2	1
x7	2	3

Compute the sum of squared error (SSE) to evaluate K-mean clustering result with the parameters below.

K:2

Initial centroids: centroid1 = (4,3) and centroid2 = (9,7)

Distance measure: Euclidean distance

7.4167

**Question 14****2 / 2 pts**

Select different aspects of cluster validation.

- Determining the clustering tendency of a set of data
- To improve the complexity of the algorithm
- Determining the 'correct' number of clusters
- None of the above

correct these are some of the aspects of cluster validation!

**Question 15****2 / 2 pts**

Select qualities of clusters produced by a good clustering algorithm.

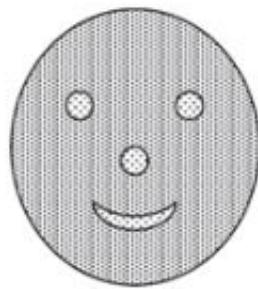
- Intra-cluster distances are minimized
- Inter-cluster distances are maximized
- Number of clusters produced
- A and B are correct

good cluster have strong cohesion within the cluster and maximum distance between them.

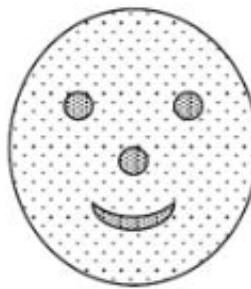
### Question 16

2 / 2 pts

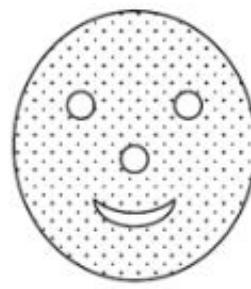
Review the image below.



(a)



(b)



(c)



(d)

**Darkness or number of dots represent density. Lines are used only to distinguish regions and do not represent points.**

If we want to find the patterns represented by the nose, eyes, and mouth using single linkage hierarchical clustering, select all figures that be well-clustered?

(a)

(b)

(c)

(d)

- None of above

**Question 17****2 / 2 pts**

Select all that are not true about DBSCAN.

- it has trouble when the clusters have widely varying densities
- it is very sensitive to noise and to the size of the data
- DBSCAN has a complexity of order  $O(n^3)$
- it can be expensive when the computation of nearest neighbors requires computing all pairwise proximities
- A and B are correct
- B and C are correct

DBSCAN is relatively resistant to noise and can handle clusters of arbitrary shapes and sizes. It has complexity order  $O(n \log n)$
- A, C and D are correct

**Question 18****2 / 2 pts**

### Select the true statement for Nearest Neighbor classification



Although Nearest Neighbor classification has the data dimensionality issue, it can be solved by the scaling technique



k-NN classifier is a typical lazy learner because it spend so much time building the model even if the 'k' is very small



Drawing a circle is a mandatory task to find neighbor classes regardless of the number of 'K'



Determining the optimal value of 'K' is important for k-NN classifier performance.



None of above

### Question 19

2 / 2 pts

Review the table below.

Rainy	Sunshine		Wind		Atmosphere		Temperature	
	Yes	No	Yes	No	High	Low	High	Low
Yes	2	6	6	2	8	0	5	3
No	8	4	2	10	2	10	6	6

Is it raining if the weather condition is *sunshine, no wind, high atmosphere, and low temperature* (Hint: you should use Naïve Bayes)?

Yes

No

## Question 20

2 / 2 pts

Review the table below.

A	B	C	Class
F	F	F	Y
F	F	T	N
F	T	T	N
F	T	T	N
F	F	T	Y
T	F	T	Y
T	F	T	N
T	F	T	N
T	T	T	Y
T	F	T	Y

Select the correct prediction class for a test sample (A=F, B=T, C=F) using the naïve Bayes approach.

---

'Y'

---

'N'

---

Anyone

---

Quiz Score: **40** out of 40

# Final Exam

**Due** Dec 9 at 1:15am      **Points** 40      **Questions** 20

**Available** Dec 6 at 12am - Dec 9 at 2am 3 days

**Time Limit** 75 Minutes

## Instructions

No collaboration

75 mins

## Attempt History

	<b>Attempt</b>	<b>Time</b>	<b>Score</b>
LATEST	<a href="#">Attempt 1</a>	43 minutes	38 out of 40

ⓘ Correct answers are hidden.

Score for this quiz: **38** out of 40

Submitted Dec 8 at 5:27pm

This attempt took 43 minutes.

### Question 1

2 / 2 pts

Transaction ID Items Bought

- 1 {Milk, Beer, Diapers}
- 2 {Bread, Butter, Milk}
- 3 {Milk, Diapers, Cookies}
- 4 {Bread, Butter, Cookies}

- 5 {Beer, Cookies, Diapers}
- 6 {Milk, Diapers, Bread, Butter}
- 7 {Bread, Butter, Diapers}
- 8 {Beer, Diapers}
- 9 {Milk, Diapers, Bread, Butter}
- 10 {Beer, Cookies}

What is the confidence of the rule Beer  $\rightarrow$  Cookies

0.5

## Question 2

2 / 2 pts

Transaction ID Items Bought

- 1 {Milk, Beer, Diapers}
- 2 {Bread, Butter, Milk}
- 3 {Milk, Diapers, Cookies}
- 4 {Bread, Butter, Cookies}
- 5 {Beer, Cookies, Diapers}
- 6 {Milk, Diapers, Bread, Butter}

- 7      {Bread, Butter, Diapers}
- 8      {Beer, Diapers}
- 9      {Milk, Diapers, Bread, Butter}
- 10     {Beer, Cookies}

What is the maximum number of size 3-itemsets that can be derived from this data set. Note that the itemsets derived may not be in the dataset.

- 
- 20
- 
- 35
- 
- 40
- 
- 15
- 
- None of the above.

### Question 3

2 / 2 pts

Example of market basket transactions.

Transaction ID Items Bought

- 1      {a, b, d, e}
- 2      {b, c, d}
- 3      {a, b, d, e}

- 4      {a, c, d, e}
- 5      {b, c, d, e}
- 6      {b, d, e}
- 7      {c, d}
- 8      {a, b, c}
- 9      {a, d, e}
- 10     {b, d}

If the minimum support is set at 40%, how many frequent 3-itemsets will be found? Note that if a 3 itemset is a subset of a larger itemset, it counts as one occurrence.

2

**Question 4****2 / 2 pts**

Transaction ID Items Bought

- 1      {Milk, Beer, Diapers}
- 2      {Bread, Butter, Milk}
- 3      {Milk, Diapers, Cookies}
- 4      {Bread, Butter, Cookies}

- 5      {Beer, Cookies, Diapers}
- 6      {Milk, Diapers, Bread, Butter}
- 7      {Bread, Butter, Diapers}
- 8      {Beer, Diapers}
- 9      {Milk, Diapers, Bread, Butter}
- 10     {Beer, Cookies}

Find all itemsets (of size 3 or larger) that has the largest support.

- 
- {Milk, Beer, Diapers}
- 
- {Bread, Butter, Milk}
- 
- {Bread, Butter, Diapers}
- 
- {Milk, Diapers, Bread}
- 
- None of the above.
- 
- All of the above.

### Question 5

2 / 2 pts

Example of market basket transactions.

Transaction ID Items Bought

- 1      {a, b, d, e}

- 2      {b, c, d}
- 3      {a, b, d, e}
- 4      {a, c, d, e}
- 5      {b, c, d, e}
- 6      {b, d, e}
- 7      {c, d}
- 8      {a, b, c}
- 9      {a, d, e}
- 10     {b, d}

Compute the confidence for the association rules  $\{a, e\} \rightarrow \{b\}$ .

0.5

### Question 6

2 / 2 pts

Usage of deep learning is predicated with availability of large amounts of data. Is the data volume the only determining factor to check whether we can use deep learning ?

Yes, lower volumes lead to overfitting



No, balance among classes in the input data is the only determining factor  
volume does not matter



No, balance among classes in the input data is a co-factor along with the data volume



None of the above

### Question 7

2 / 2 pts

Select all that is true about Restricted Boltzman Machine (RBM)

Can be used for recommendation systems

RBM training is often probabilistic

RBM is a kind of Deep Belief Network

None of the above.

### Question 8

2 / 2 pts

Which of the following is an application of autoencoders?

feature learning.

dimensionality reduction.

Rule mining

all of the above.

**Question 9****2 / 2 pts**

Increase in size of a convolutional hidden layer would not necessarily increase the performance of a convolutional network.

---

True

---

False

**Question 10****2 / 2 pts**

Select the true statements about neural network (NN) complexity.

---

CNN is always more complex than DNN

---

Adding more hidden layers increases complexity of a NN

---

An Autoencoder is always more complex than a CNN

---



Complexity metric defined for a NN may not make sense for other classification engines

**Question 11****2 / 2 pts**

Review the following table.

Data	X	Y
A	1	2
B	2	2

C	2	3
D	8	7
E	8	8
F	8	2

Select the correct number of core points and DBSCAN clustering result with the parameters below.

Eps: 5

Minimum samples (MinPts):3

Distance measurement: Euclidean distance

- 5, cluster1: {A,B,C}, cluster2: {D,E,F}
- 6, cluster1: {A,B}, cluster2: {C,D,E}, cluster3: {F}
- 4, cluster1: {A,B,C}, cluster2: {D, E, F}
- 5, cluster1: {A}, cluster2: {B}, cluster3: {C}, cluster4: {D}, cluster5: {E}
- None of above

### Question 12

2 / 2 pts

Select qualities of clusters produced by a good clustering algorithm.

- Intra-cluster distances are minimized
- Inter-cluster distances are maximized
- Number of clusters produced
- A and B are correct

good cluster have strong cohesion within the cluster and maximum distance between them.

**Question 13****2 / 2 pts**

What is the purpose of cluster analysis?

- To avoid finding patterns in noise
- To compare clustering algorithms
- To compare two sets of clusters
- all of the above
- None of the above

we do cluster validity to avoid clustering noise and find a suitable algorithim for our data

**Question 14****2 / 2 pts**

Review the following table.

Data	X	Y
x1	1	1
x2	6	6
x3	7	7
x4	3	2
x5	6	8
x6	2	1
x7	2	3

Select the correct pair of final centroid and clustering set using K-mean clustering with the parameters below.

K:2

Initial centroids: centroid1 = (4,3) and centroid2 = (9,7)

Distance measure: Euclidean distance



Cluster 1: (x1, x4, x6, x7) with centroid=(2, 1.75), Cluster 2: (x2, x3, x5) with centroid=(6.33, 7)



Cluster 1: (x1, x4, x6, x7) with centroid=(1.75, 2), Cluster 2: (x2, x3, x5) with centroid=(6.33, 7)



Cluster 1: (x1, x4, x6, x7) with centroid=(2, 1.75), Cluster 2: (x2, x3, x5) with centroid=(7, 6.33)



Cluster 1: (x1, x2, x4, x6, x7) with centroid=(2, 1.75), Cluster 2: (x3, x5) with centroid=(6.33, 7)



Cluster 1:  $(x_1, x_2, x_4, x_6, x_7)$  with centroid=(2, 1.75), Cluster 2:  $(x_3, x_5)$  with centroid=(7, 6.33)

**Question 15****2 / 2 pts**

Review the following table.

Data	X	Y
x1	1	1
x2	6	6
x3	7	7
x4	3	2
x5	6	8
x6	2	1
x7	2	3

Compute the sum of squared error (SSE) to evaluate K-mean clustering result with the parameters below.

K:2

Initial centroids: centroid1 = (4,3) and centroid2 = (9,7)

Distance measure: Euclidean distance

7.4167

**Question 16****2 / 2 pts**

Select all statements that are true for Hierarchical clustering.

Although it requires high storage, high computing is not required

Robust to noise

Robust to outliers

It is not easy to decide on the number of cluster because the complicated dendrogram disturbs this decision

None of above

Incorrect

### Question 17

0 / 2 pts

Review the following table.

Data	X	Y
x1	1	1
x2	6	6
x3	7	7
x4	3	2
x5	6	8
x6	2	1
x7	2	3

Compute the updated centroids after the first iteration using K-mean clustering with the parameters below.

K:2

Initial centroids: centroid1 = (4,3) and centroid2 = (9,7)

Distance measure: Euclidean

Type your answer in the same format as "centroid1 = (x1,y1) and centroid2 = (x2,y2)"

Each co-ordinate should be in the format A.BC, where A, B and C are integers. If you have 4 as one of the co-ordinates you should write 4.00. Please be mindful of the spaces and other formats.

centroid1 = (2.00,1.75) and centroid2 = (6.33,7.00)

### Question 18

2 / 2 pts

Review the table below.

Instance	X-Axis	Class
i	-3	-
ii	-2	-
iii	-1	+
iv	0	+
v	1	+
vi	2	-

Select all the support vector instances using Non-linear SVM

Feature Map ( $\{R^1 \rightarrow R^2\}$ )

i

ii iii iv v vi**Question 19****2 / 2 pts**

Review the table below.

A	B	C	Class
F	F	F	Y
F	F	T	N
F	T	T	N
F	T	T	N
F	F	T	Y
T	F	T	Y
T	F	T	N
T	F	T	N
T	T	T	Y
T	F	T	Y

Select the correct prediction class for a test sample (A=F, B=T, C=F) using the naïve Bayes approach.

 'Y'

'N' Anyone**Question 20****2 / 2 pts**

Review the table below.

Instance	X-Axis	Class
i	-3	-
ii	-2	-
iii	-1	+
iv	0	+
v	1	+
vi	2	-

Select a correct hyper-plane equation of the data below using Non-linear SVM.

**Feature Map ( $\{R^1 \rightarrow R^2\}$ )** Y=1 Y=1.5 Y=2 Y=2.5 None of above

Quiz Score: **38** out of 40

# Final Exam

**Due** Dec 9 at 1:15am      **Points** 40      **Questions** 20

**Available** Dec 6 at 12am - Dec 9 at 2am 3 days

**Time Limit** 75 Minutes

## Instructions

No collaboration

75 mins

## Attempt History

	<b>Attempt</b>	<b>Time</b>	<b>Score</b>
LATEST	<a href="#">Attempt 1</a>	43 minutes	38 out of 40

! Correct answers are hidden.

Score for this quiz: 38 out of 40

Submitted Dec 8 at 5:27pm

This attempt took 43 minutes.

### Question 1

2 / 2 pts

Transaction ID Items Bought

- 1 {Milk, Beer, Diapers}
- 2 {Bread, Butter, Milk}
- 3 {Milk, Diapers, Cookies}
- 4 {Bread, Butter, Cookies}

- 5      {Beer, Cookies, Diapers}
- 6      {Milk, Diapers, Bread, Butter}
- 7      {Bread, Butter, Diapers}
- 8      {Beer, Diapers}
- 9      {Milk, Diapers, Bread, Butter}
- 10     {Beer, Cookies}

What is the confidence of the rule Beer  $\rightarrow$  Cookies

0.5

## Question 2

2 / 2 pts

Transaction ID Items Bought

- 1      {Milk, Beer, Diapers}
- 2      {Bread, Butter, Milk}
- 3      {Milk, Diapers, Cookies}
- 4      {Bread, Butter, Cookies}
- 5      {Beer, Cookies, Diapers}
- 6      {Milk, Diapers, Bread, Butter}

- 7      {Bread, Butter, Diapers}
- 8      {Beer, Diapers}
- 9      {Milk, Diapers, Bread, Butter}
- 10     {Beer, Cookies}

What is the maximum number of size 3-itemsets that can be derived from this data set. Note that the itemsets derived may not be in the dataset.

- 
- 20
- 
- 35
- 
- 40
- 
- 15
- 
- None of the above.

### Question 3

2 / 2 pts

Example of market basket transactions.

Transaction ID Items Bought

- 1      {a, b, d, e}
- 2      {b, c, d}
- 3      {a, b, d, e}

- 4      {a, c, d, e}
- 5      {b, c, d, e}
- 6      {b, d, e}
- 7      {c, d}
- 8      {a, b, c}
- 9      {a, d, e}
- 10     {b, d}

If the minimum support is set at 40%, how many frequent 3-itemsets will be found? Note that if a 3 itemset is a subset of a larger itemset, it counts as one occurrence.

2

**Question 4****2 / 2 pts**

Transaction ID Items Bought

- 1      {Milk, Beer, Diapers}
- 2      {Bread, Butter, Milk}
- 3      {Milk, Diapers, Cookies}
- 4      {Bread, Butter, Cookies}

- 5      {Beer, Cookies, Diapers}
- 6      {Milk, Diapers, Bread, Butter}
- 7      {Bread, Butter, Diapers}
- 8      {Beer, Diapers}
- 9      {Milk, Diapers, Bread, Butter}
- 10     {Beer, Cookies}

Find all itemsets (of size 3 or larger) that has the largest support.

- 
- {Milk, Beer, Diapers}
- 
- {Bread, Butter, Milk}
- 
- {Bread, Butter, Diapers}
- 
- {Milk, Diapers, Bread}
- 
- None of the above.
- 
- All of the above.

### Question 5

2 / 2 pts

Example of market basket transactions.

Transaction ID Items Bought

- 1      {a, b, d, e}

- 2      {b, c, d}
- 3      {a, b, d, e}
- 4      {a, c, d, e}
- 5      {b, c, d, e}
- 6      {b, d, e}
- 7      {c, d}
- 8      {a, b, c}
- 9      {a, d, e}
- 10     {b, d}

Compute the confidence for the association rules  $\{a, e\} \rightarrow \{b\}$ .

0.5

### Question 6

2 / 2 pts

Usage of deep learning is predicated with availability of large amounts of data. Is the data volume the only determining factor to check whether we can use deep learning ?

Yes, lower volumes lead to overfitting



No, balance among classes in the input data is the only determining factor  
volume does not matter



No, balance among classes in the input data is a co-factor along with the data volume



None of the above

### Question 7

2 / 2 pts

Select all that is true about Restricted Boltzman Machine (RBM)

Can be used for recommendation systems

RBM training is often probabilistic

RBM is a kind of Deep Belief Network

None of the above.

### Question 8

2 / 2 pts

Which of the following is an application of autoencoders?

feature learning.

dimensionality reduction.

Rule mining

all of the above.

**Question 9****2 / 2 pts**

Increase in size of a convolutional hidden layer would not necessarily increase the performance of a convolutional network.

---

True

---

False

**Question 10****2 / 2 pts**

Select the true statements about neural network (NN) complexity.

---

CNN is always more complex than DNN

---

Adding more hidden layers increases complexity of a NN

---

An Autoencoder is always more complex than a CNN

---



Complexity metric defined for a NN may not make sense for other classification engines

**Question 11****2 / 2 pts**

Review the following table.

Data	X	Y
A	1	2
B	2	2

C	2	3
D	8	7
E	8	8
F	8	2

Select the correct number of core points and DBSCAN clustering result with the parameters below.

Eps: 5

Minimum samples (MinPts):3

Distance measurement: Euclidean distance

- 5, cluster1: {A,B,C}, cluster2: {D,E,F}
- 6, cluster1: {A,B}, cluster2: {C,D,E}, cluster3: {F}
- 4, cluster1: {A,B,C}, cluster2: {D, E, F}
- 5, cluster1: {A}, cluster2: {B}, cluster3: {C}, cluster4: {D}, cluster5: {E}
- None of above

### Question 12

2 / 2 pts

Select qualities of clusters produced by a good clustering algorithm.

- Intra-cluster distances are minimized
- Inter-cluster distances are maximized
- Number of clusters produced
- A and B are correct

good cluster have strong cohesion within the cluster and maximum distance between them.

**Question 13****2 / 2 pts**

What is the purpose of cluster analysis?

- To avoid finding patterns in noise
- To compare clustering algorithms
- To compare two sets of clusters
- all of the above
- None of the above

we do cluster validity to avoid clustering noise and find a suitable algorithim for our data

**Question 14****2 / 2 pts**

Review the following table.

Data	X	Y
x1	1	1
x2	6	6
x3	7	7
x4	3	2
x5	6	8
x6	2	1
x7	2	3

Select the correct pair of final centroid and clustering set using K-mean clustering with the parameters below.

K:2

Initial centroids: centroid1 = (4,3) and centroid2 = (9,7)

Distance measure: Euclidean distance



Cluster 1: (x1, x4, x6, x7) with centroid=(2, 1.75), Cluster 2: (x2, x3, x5) with centroid=(6.33, 7)



Cluster 1: (x1, x4, x6, x7) with centroid=(1.75, 2), Cluster 2: (x2, x3, x5) with centroid=(6.33, 7)



Cluster 1: (x1, x4, x6, x7) with centroid=(2, 1.75), Cluster 2: (x2, x3, x5) with centroid=(7, 6.33)



Cluster 1: (x1, x2, x4, x6, x7) with centroid=(2, 1.75), Cluster 2: (x3, x5) with centroid=(6.33, 7)



Cluster 1:  $(x_1, x_2, x_4, x_6, x_7)$  with centroid=(2, 1.75), Cluster 2:  $(x_3, x_5)$  with centroid=(7, 6.33)

**Question 15****2 / 2 pts**

Review the following table.

Data	X	Y
x1	1	1
x2	6	6
x3	7	7
x4	3	2
x5	6	8
x6	2	1
x7	2	3

Compute the sum of squared error (SSE) to evaluate K-mean clustering result with the parameters below.

K:2

Initial centroids: centroid1 = (4,3) and centroid2 = (9,7)

Distance measure: Euclidean distance

7.4167

**Question 16****2 / 2 pts**

Select all statements that are true for Hierarchical clustering.

Although it requires high storage, high computing is not required

Robust to noise

Robust to outliers

It is not easy to decide on the number of cluster because the complicated dendrogram disturbs this decision

None of above

Incorrect

### Question 17

0 / 2 pts

Review the following table.

Data	X	Y
x1	1	1
x2	6	6
x3	7	7
x4	3	2
x5	6	8
x6	2	1
x7	2	3

Compute the updated centroids after the first iteration using K-mean clustering with the parameters below.

K:2

Initial centroids: centroid1 = (4,3) and centroid2 = (9,7)

Distance measure: Euclidean

Type your answer in the same format as "centroid1 = (x1,y1) and centroid2 = (x2,y2)"

Each co-ordinate should be in the format A.BC, where A, B and C are integers. If you have 4 as one of the co-ordinates you should write 4.00. Please be mindful of the spaces and other formats.

centroid1 = (2.00,1.75) and centroid2 = (6.33,7.00)

### Question 18

2 / 2 pts

Review the table below.

Instance	X-Axis	Class
i	-3	-
ii	-2	-
iii	-1	+
iv	0	+
v	1	+
vi	2	-

Select all the support vector instances using Non-linear SVM

Feature Map ( $\{R^1 \rightarrow R^2\}$ )

i

ii iii iv v vi**Question 19****2 / 2 pts**

Review the table below.

A	B	C	Class
F	F	F	Y
F	F	T	N
F	T	T	N
F	T	T	N
F	F	T	Y
T	F	T	Y
T	F	T	N
T	F	T	N
T	T	T	Y
T	F	T	Y

Select the correct prediction class for a test sample (A=F, B=T, C=F) using the naïve Bayes approach.

 'Y'

'N' Anyone**Question 20****2 / 2 pts**

Review the table below.

Instance	X-Axis	Class
i	-3	-
ii	-2	-
iii	-1	+
iv	0	+
v	1	+
vi	2	-

Select a correct hyper-plane equation of the data below using Non-linear SVM.

**Feature Map ( $\{R^1 \rightarrow R^2\}$ )** Y=1 Y=1.5 Y=2 Y=2.5 None of above

Quiz Score: **38** out of 40

# Final Exam

**Due** Dec 9 at 1:15am      **Points** 40      **Questions** 20

**Available** Dec 6 at 12am - Dec 9 at 2am 3 days

**Time Limit** 75 Minutes

## Instructions

No collaboration

75 mins

## Attempt History

	Attempt	Time	Score
LATEST	<a href="#"><u>Attempt 1</u></a>	46 minutes	38 out of 40

! Correct answers are hidden.

Score for this quiz: **38** out of 40

Submitted Dec 8 at 4:43pm

This attempt took 46 minutes.

Question 1	2 / 2 pts
Example of market basket transactions.	
Transaction ID Items Bought	
1	{a, b, d, e}
2	{b, c, d}

- |    |              |
|----|--------------|
| 3  | {a, b, d, e} |
| 4  | {a, c, d, e} |
| 5  | {b, c, d, e} |
| 6  | {b, d, e}    |
| 7  | {c, d}       |
| 8  | {a, b, c}    |
| 9  | {a, d, e}    |
| 10 | {b, d}       |

If the minimum support is set at 40%, how many frequent 3-itemsets will be found? Note that if a 3 itemset is a subset of a larger itemset, it counts as one occurrence.

2

**Question 2****2 / 2 pts**

Select the reason why Apriori pruning in the search for frequent itemsets works.

- support count is monotonic with respect to itemsets.
- we search in transaction ID order.
- support count diverges as we add to the itemset.

- support count is anti-monotonic with respect to itemsets.
- it is an excellent heuristic, but it does not work 100% of the time.

**Question 3****2 / 2 pts**

Transaction ID Items Bought

- 1 {Milk, Beer, Diapers}
- 2 {Bread, Butter, Milk}
- 3 {Milk, Diapers, Cookies}
- 4 {Bread, Butter, Cookies}
- 5 {Beer, Cookies, Diapers}
- 6 {Milk, Diapers, Bread, Butter}
- 7 {Bread, Butter, Diapers}
- 8 {Beer, Diapers}
- 9 {Milk, Diapers, Bread, Butter}
- 10 {Beer, Cookies}

What is the maximum size of frequent itemsets that can be extracted  
(assuming minsup > 0.1)?

10 9 4 3**Question 4****2 / 2 pts**

Transaction ID Items Bought

- 1 {Milk, Beer, Diapers}
- 2 {Bread, Butter, Milk}
- 3 {Milk, Diapers, Cookies}
- 4 {Bread, Butter, Cookies}
- 5 {Beer, Cookies, Diapers}
- 6 {Milk, Diapers, Bread, Butter}
- 7 {Bread, Butter, Diapers}
- 8 {Beer, Diapers}
- 9 {Milk, Diapers, Bread, Butter}
- 10 {Beer, Cookies}

Find all itemsets (of size 3 or larger) that has the largest support.

{Milk, Beer, Diapers}

{Bread, Butter, Milk}

{Bread, Butter, Diapers}

{Milk, Diapers, Bread}

None of the above.

All of the above.

**Question 5****2 / 2 pts**

Which of the above statements are true for any A, B, and C?

- If  $A \rightarrow B$  then  $B \rightarrow A$ .
- If  $A \rightarrow B$  and  $B \rightarrow C$  then  $A \rightarrow C$ .
- If  $A \rightarrow C$  then  $A \cup B \rightarrow C$ .
- If  $A \cup B \rightarrow C$  then  $A \rightarrow C$ .
- I & II
- I, II, & III
- I, II, & IV
- II & III
- none

**Question 6****2 / 2 pts**

The number of neurons in the output layer should match the number of classes (Where the number of classes is greater than 2) in a supervised learning task.

- True
- False

**Question 7****2 / 2 pts**

Increase in size of a convolutional hidden layer would not necessarily increase the performance of a convolutional network.

True

False

### Question 8

2 / 2 pts

Is an autoencoder the same as performing Principal Component Analysis?

Not at all

No but conceptually they can be used for the same purpose

Autoencoders can do dimensionality reduction, but it is non-linear.

Yes they have the same purpose.

### Question 9

2 / 2 pts

Usage of deep learning is predicated with availability of large amounts of data. Is the data volume the only determining factor to check whether we can use deep learning ?

Yes, lower volumes lead to overfitting

No, balance among classes in the input data is the only determining factor volume does not matter



No, balance among classes in the input data is a co-factor along with the data volume



None of the above

### Question 10

2 / 2 pts

Which of the following is an application of autoencoders?

feature learning.

dimensionality reduction.

Rule mining

all of the above.

### Question 11

2 / 2 pts

Select all statements that are true for Hierarchical clustering.

Although it requires high storage, high computing is not required

Robust to noise

Robust to outliers

It is not easy to decide on the number of cluster because the complicated dendrogram disturbs this decision

- None of above

### Question 12

2 / 2 pts

Review the table below.

Cluster	Tennis	Golf	Baseball	Basketball	Football	Hockey	Total	Entropy	Purity
#1	1	1	0	11	4	676	693		
#2	27	89	333	827	253	33	1562		
#3	326	465	8	105	16	29	949		
Total	354	555	341	943	273	738	3204		

Select the collect purity set of the confusion matrix

- Cluster #1: 0.98, Cluster #2: 0.53, Cluster #3: 0.49, Total: 0.61

- Cluster #1: 0.53, Cluster #2: 0.98, Cluster #3: 0.61, Total: 0.49

- Cluster #1: 0.98, Cluster #2: 0.49, Cluster #3: 0.53, Total: 0.61

- Cluster #1: 0.53, Cluster #2: 0.49, Cluster #3: 0.61, Total: 0.98

- None of above

### Question 13

2 / 2 pts

Select all that are not true about DBSCAN.

- It has trouble when the clusters have widely varying densities

- it is very sensitive to noise and to the size of the data
- DBSCAN has a complexity of order  $O(n^3)$
- it can be expensive when the computation of nearest neighbors requires computing all pairwise proximities
- A and B are correct
- B and C are correct
  - DBSCAN is relatively resistant to noise and can handle clusters of arbitrary shapes and sizes. it has complexity order  $O(n \log n)$
- A, C and D are correct

**Question 14****2 / 2 pts**

Review the following table.

Data	X	Y
x1	1	1
x2	6	6
x3	7	7
x4	3	2
x5	6	8
x6	2	1
x7	2	3

Select the correct pair of final centroid and clustering set using K-mean clustering with the parameters below.

K:2

Initial centroids: centroid1 = (4,3) and centroid2 = (9,7)

Distance measure: Euclidean distance



Cluster 1: (x1, x4, x6, x7) with centroid=(2, 1.75), Cluster 2: (x2, x3, x5) with centroid=(6.33, 7)



Cluster 1: (x1, x4, x6, x7) with centroid=(1.75, 2), Cluster 2: (x2, x3, x5) with centroid=(6.33, 7)



Cluster 1: (x1, x4, x6, x7) with centroid=(2, 1.75), Cluster 2: (x2, x3, x5) with centroid=(7, 6.33)



Cluster 1: (x1, x2, x4, x6, x7) with centroid=(2, 1.75), Cluster 2: (x3, x5) with centroid=(6.33, 7)



Cluster 1: (x1, x2, x4, x6, x7) with centroid=(2, 1.75), Cluster 2: (x3, x5) with centroid=(7, 6.33)

### Question 15

2 / 2 pts

Review the following table.

Data	X	Y
x1	1	1
x2	6	6
x3	7	7
x4	3	2
x5	6	8
x6	2	1
x7	2	3

Compute the sum of squared error (SSE) to evaluate K-mean clustering result with the parameters below.

K:2

Initial centroids: centroid1 = (4,3) and centroid2 = (9,7)

Distance measure: Euclidean distance

7.4167

Incorrect

### Question 16

0 / 2 pts

Review the following table.

Data	X	Y
x1	1	1
x2	6	6
x3	7	7
x4	3	2
x5	6	8
x6	2	1
x7	2	3

Compute the updated centroids after the first iteration using K-mean clustering with the parameters below.

K:2

Initial centroids: centroid1 = (4,3) and centroid2 = (9,7)

Distance measure: Euclidean

Type your answer in the same format as "centroid1 = (x1,y1) and centroid2 = (x2,y2)"

Each co-ordinate should be in the format A.BC, where A, B and C are integers. If you have 4 as one of the co-ordinates you should write 4.00. Please be mindful of the spaces and other formats.

centroid1 = (2.00,1.75) and centroid2 = (6.33,7.00)

### Question 17

2 / 2 pts

Review the following table.

	A	B	C	D	E
A	0	9	3	6	11
B	9	0	7	5	10
C	3	7	0	9	2
D	6	5	9	0	8
E	11	10	2	8	0

Select the correct DBSCAN clustering result with the parameters below

epsilon = 5, Minpts = 2

- cluster1: {A, C, E}, cluster2: {B,D}
- cluster1: {A}, cluster2: {B}, cluster3: {D}, and cluster4: {C, E}
- cluster1: {A, B}, cluster2: {C}, and cluster3: {D}, cluster4: {E}
- cluster1: {A}, cluster2: {B, E}, cluster3: {C}, and cluster4: {D}
- None of above

### Question 18

2 / 2 pts

Review the table below.

A	B	C	Class
F	F	F	Y
F	F	T	N
F	T	T	N
F	T	T	N
F	F	T	Y
T	F	T	Y
T	F	T	N
T	F	T	N
T	T	T	Y
T	F	T	Y

Select the correct prediction class for a test sample (A=F, B=T, C=F) using the naïve Bayes approach.

'Y'

'N'

Anyone

### Question 19

2 / 2 pts

Select the true statement for Nearest Neighbor classification

Although Nearest Neighbor classification has the data dimensionality issue, it can be solved by the scaling technique

k-NN classifier is a typical lazy learner because it spend so much time building the model even if the 'k' is very small

Drawing a circle is a mandatory task to find neighbor classes regardless of the number of 'K'

Determining the optimal value of 'K' is important for k-NN classifier performance.

None of above

### Question 20

2 / 2 pts

Review the table below.

Instance	X-Axis	Class
i	-3	-
ii	-2	-
iii	-1	+
iv	0	+
v	1	+
vi	2	-

Select all the support vector instances using Non-linear SVM

Feature Map ( $\{R^1 \rightarrow R^2\}$ )

 i

<input checked="" type="checkbox"/> ii
<input checked="" type="checkbox"/> iii
<input type="checkbox"/> iv
<input checked="" type="checkbox"/> v
<input checked="" type="checkbox"/> vi

Quiz Score: **38** out of 40

# Final Exam

**Due Dec 9 at 1:15am      Points 40      Questions 20**

**Available** Dec 6 at 12am - Dec 9 at 2am 3 days

**Time Limit** 75 Minutes

## Instructions

No collaboration

75 mins

## Attempt History

	<b>Attempt</b>	<b>Time</b>	<b>Score</b>
<b>LATEST</b>	<a href="#"><u>Attempt 1</u></a>	41 minutes	40 out of 40

! Correct answers are hidden.

Score for this quiz: **40** out of 40

Submitted Dec 8 at 6:17pm

This attempt took 41 minutes.

<b>Question 1</b>	<b>2 / 2 pts</b>
Apriori pruning based on support is a greedy strategy but is not optimal	
<input type="radio"/> True	
<input checked="" type="radio"/> False	

**Question 2****2 / 2 pts**

Which of the above statements are true for any A, B, and C?

- If  $A \rightarrow B$  then  $B \rightarrow A$ .
- If  $A \rightarrow B$  and  $B \rightarrow C$  then  $A \rightarrow C$ .
- If  $A \rightarrow C$  then  $A \cup B \rightarrow C$ .
- If  $A \cup B \rightarrow C$  then  $A \rightarrow C$ .
- I & II
- I, II, & III
- I, II, & IV
- II & III
- none

**Question 3****2 / 2 pts**

Example of market basket transactions.

Transaction ID Items Bought

- |    |              |
|----|--------------|
| 1  | {a, b, d, e} |
| 2  | {b, c, d}    |
| 3  | {a, b, d, e} |
| 4  | {a, c, d, e} |
| 5  | {b, c, d, e} |
| 6  | {b, d, e}    |
| 7  | {c, d}       |
| 8  | {a, b, c}    |
| 9  | {a, d, e}    |
| 10 | {b, d}       |

Compute the confidence for the association rules {a, e} -> {b}.

0.5

#### Question 4

2 / 2 pts

Example of market basket transactions.

**Transaction ID Items Bought**

1 {a, b, d, e}

2 {b, c, d}

3 {a, b, d, e}

4 {a, c, d, e}

5 {b, c, d, e}

6 {b, d, e}

7 {c, d}

8 {a, b, c}

9 {a, d, e}

10 {b, d}

If the minimum support is set at 40%, how many frequent 3-itemsets will be found? Note that if a 3 itemset is a subset of a larger itemset, it counts as one occurrence.

2

**Question 5****2 / 2 pts**

## Example of market basket transactions.

### Transaction ID Items Bought

1 {a, b, d, e}

2 {b, c, d}

3 {a, b, d, e}

4 {a, c, d, e}

5 {b, c, d, e}

6 {b, d, e}

7 {c, d}

8 {a, b, c}

9 {a, d, e}

10 {b, d}

Compute the support for itemset {b,c,d}.

0.2

**Question 6****2 / 2 pts**

Which of the following is an application of autoencoders?

- 
- feature learning.
- 
- dimensionality reduction.
- 
- Rule mining
- 
- all of the above.

**Question 7****2 / 2 pts**

What is an autoencoder?

- 
- a neural network that copies its input to its output.
- 
- a neural network that codes itself.
- 
- 
- a neural network that maps an output to an input through a hidden layer.
- 
- 
- a neural network that is trained to attempt to copy its input to its output.

**Question 8****2 / 2 pts**

The number of neurons in the output layer should match the number of classes (Where the number of classes is greater than 2) in a supervised learning task.

 True False**Question 9****2 / 2 pts**

Select all that is true about Restricted Boltzman Machine (RBM)

 Can be used for recommendation systems RBM training is often probabilistic RBM is a kind of Deep Belief Network None of the above.**Question 10****2 / 2 pts**

Consider a movie recommendation application, where an user can log in and select movies they like based on their previous preferences.

Which neural network architecture would be suitable to complete this task

- Fully-Connected Neural Network.
- Convolutional Neural Network.
- Recurrent Neural Network.
- Restricted Boltzmann Machine.

### Question 11

2 / 2 pts

Review the following table.

Data	X	Y
x1	1	1
x2	6	6
x3	7	7
x4	3	2
x5	6	8
x6	2	1
x7	2	3

Compute the average silhouette score to evaluate K-mean clustering result with the parameters below.

K:2

Initial centroids: centroid1 = (4,3) and centroid2 = (9,7)

Distance measure: Euclidean distance

0.752

### Question 12

2 / 2 pts

Review the following table.

Data	X	Y
A	1	2
B	2	2
C	2	3
D	8	7
E	8	8
F	8	2

Select the correct number of core points and DBSCAN clustering result with the parameters below.

Eps: 5

Minimum samples (MinPts):3

Distance measurement: Euclidean distance

- 5, cluster1: {A,B,C}, cluster2: {D,E,F}
- 6, cluster1: {A,B}, cluster2: {C,D,E}, cluster3: {F}
- 4, cluster1: {A,B,C}, cluster2: {D, E, F}



5, cluster1: {A}, cluster2: {B}, cluster3: {C}, cluster4: {D}, cluster5: {E}



None of above

### Question 13

2 / 2 pts

Review the following table.

Data	X	Y
x1	1	1
x2	6	6
x3	7	7
x4	3	2
x5	6	8
x6	2	1
x7	2	3

Compute the sum of squared error (SSE) to evaluate K-mean clustering result with the parameters below.

K:2

Initial centroids: centroid1 = (4,3) and centroid2 = (9,7)

Distance measure: Euclidean distance

7.4167

**Question 14****2 / 2 pts**

Select different aspects of cluster validation.

- Determining the clustering tendency of a set of data
- To improve the complexity of the algorithm
- Determining the 'correct' number of clusters
- None of the above

correct these are some of the aspects of cluster validation!

**Question 15****2 / 2 pts**

Select qualities of clusters produced by a good clustering algorithm.

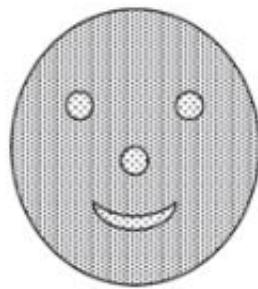
- Intra-cluster distances are minimized
- Inter-cluster distances are maximized
- Number of clusters produced
- A and B are correct

good cluster have strong cohesion within the cluster and maximum distance between them.

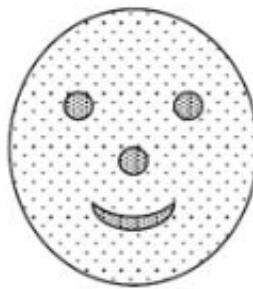
### Question 16

2 / 2 pts

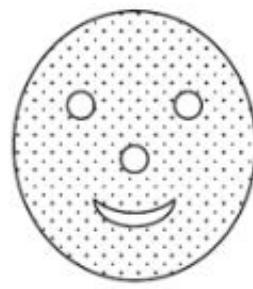
Review the image below.



(a)



(b)



(c)



(d)

**Darkness or number of dots represent density. Lines are used only to distinguish regions and do not represent points.**

If we want to find the patterns represented by the nose, eyes, and mouth using single linkage hierarchical clustering, select all figures that be well-clustered?

(a)

(b)

(c)

(d)

- None of above

**Question 17****2 / 2 pts**

Select all that are not true about DBSCAN.

- it has trouble when the clusters have widely varying densities
- it is very sensitive to noise and to the size of the data
- DBSCAN has a complexity of order  $O(n^3)$
- it can be expensive when the computation of nearest neighbors requires computing all pairwise proximities
- A and B are correct
- B and C are correct

DBSCAN is relatively resistant to noise and can handle clusters of arbitrary shapes and sizes. It has complexity order  $O(n \log n)$
- A, C and D are correct

**Question 18****2 / 2 pts**

### Select the true statement for Nearest Neighbor classification



Although Nearest Neighbor classification has the data dimensionality issue, it can be solved by the scaling technique



k-NN classifier is a typical lazy learner because it spend so much time building the model even if the 'k' is very small



Drawing a circle is a mandatory task to find neighbor classes regardless of the number of 'K'



Determining the optimal value of 'K' is important for k-NN classifier performance.



None of above

### Question 19

2 / 2 pts

Review the table below.

Rainy	Sunshine		Wind		Atmosphere		Temperature	
	Yes	No	Yes	No	High	Low	High	Low
Yes	2	6	6	2	8	0	5	3
No	8	4	2	10	2	10	6	6

Is it raining if the weather condition is *sunshine, no wind, high atmosphere, and low temperature* (Hint: you should use Naïve Bayes)?

Yes

No

## Question 20

2 / 2 pts

Review the table below.

A	B	C	Class
F	F	F	Y
F	F	T	N
F	T	T	N
F	T	T	N
F	F	T	Y
T	F	T	Y
T	F	T	N
T	F	T	N
T	T	T	Y
T	F	T	Y

Select the correct prediction class for a test sample (A=F, B=T, C=F) using the naïve Bayes approach.

---

'Y'

---

'N'

---

Anyone

---

Quiz Score: **40** out of 40

# Final Exam

**Due** Dec 9 at 1:15am      **Points** 40      **Questions** 20

**Available** Dec 6 at 12am - Dec 9 at 2am 3 days

**Time Limit** 75 Minutes

## Instructions

No collaboration

75 mins

## Attempt History

	<b>Attempt</b>	<b>Time</b>	<b>Score</b>
LATEST	<a href="#"><u>Attempt 1</u></a>	44 minutes	40 out of 40

! Correct answers are hidden.

Score for this quiz: **40** out of 40

Submitted Dec 8 at 6pm

This attempt took 44 minutes.

### Question 1

2 / 2 pts

Transaction ID Items Bought

- 1 {Milk, Beer, Diapers}
- 2 {Bread, Butter, Milk}
- 3 {Milk, Diapers, Cookies}
- 4 {Bread, Butter, Cookies}

- 5        {Beer, Cookies, Diapers}
- 6        {Milk, Diapers, Bread, Butter}
- 7        {Bread, Butter, Diapers}
- 8        {Beer, Diapers}
- 9        {Milk, Diapers, Bread, Butter}
- 10      {Beer, Cookies}

What is the maximum size of frequent itemsets that can be extracted  
(assuming minsup > 0.1)?

---

10

---

9

---

4

---

3

---

## Question 2

2 / 2 pts

Which of the above statements are true for any A, B, and C?

---

If A → B then B → A.

---

If A → B and B → C then A → C.

---

If A → C then A union B → C.

---

If A union B  $\rightarrow$  C then A  $\rightarrow$  C.

I & II

I, II, & III

I, II, & IV

II & III

none

### Question 3

2 / 2 pts

Example of market basket transactions.

#### Transaction ID Items Bought

1 {a, b, d, e}

2 {b, c, d}

3 {a, b, d, e}

4 {a, c, d, e}

5 {b, c, d, e}

6 {b, d, e}

7 {c, d}

8 {a, b, c}

9      {a, d, e}

10     {b, d}

Compute the confidence for the association rules {a, e} -> {b}.

0.5

#### Question 4

2 / 2 pts

Transaction ID Items Bought

1      {Milk, Beer, Diapers}

2      {Bread, Butter, Milk}

3      {Milk, Diapers, Cookies}

4      {Bread, Butter, Cookies}

5      {Beer, Cookies, Diapers}

6      {Milk, Diapers, Bread, Butter}

7      {Bread, Butter, Diapers}

8      {Beer, Diapers}

9      {Milk, Diapers, Bread, Butter}

10     {Beer, Cookies}

What is the maximum number of size 3-itemsets that can be derived from this data set. Note that the itemsets derived may not be in the dataset.

20

35

40

15

None of the above.

### Question 5

2 / 2 pts

Transaction ID Items Bought

- |   |                                |
|---|--------------------------------|
| 1 | {Milk, Beer, Diapers}          |
| 2 | {Bread, Butter, Milk}          |
| 3 | {Milk, Diapers, Cookies}       |
| 4 | {Bread, Butter, Cookies}       |
| 5 | {Beer, Cookies, Diapers}       |
| 6 | {Milk, Diapers, Bread, Butter} |
| 7 | {Bread, Butter, Diapers}       |
| 8 | {Beer, Diapers}                |

9 {Milk, Diapers, Bread, Butter}

10 {Beer, Cookies}

What is the confidence of the rule Beer -> Cookies

0.5

### Question 6

2 / 2 pts

Increase in size of a convolutional hidden layer would not necessarily increase the performance of a convolutional network.

True

False

### Question 7

2 / 2 pts

Consider a movie recommendation application, where an user can log in and select movies they like based on their previous preferences.

Which neural network architecture would be suitable to complete this task

Fully-Connected Neural Network.

Convolutional Neural Network.

Recurrent Neural Network.

Restricted Boltzmann Machine.

**Question 8****2 / 2 pts**

What is an autoencoder?

- a neural network that copies its input to its output.
- a neural network that codes itself.
- 
- a neural network that maps an output to an input through a hidden layer.
- a neural network that is trained to attempt to copy its input to its output.

**Question 9****2 / 2 pts**

Which of the following is an application of autoencoders?

- feature learning.
- dimensionality reduction.
- Rule mining
- all of the above.

**Question 10****2 / 2 pts**

Is an autoencoder the same as performing Principal Component Analysis?

- Not at all
- No but conceptually they can be used for the same purpose
- Autoencoders can do dimensionality reduction, but it is non-linear.
- Yes they have the same purpose.

**Question 11****2 / 2 pts**

Review the table below.

Cluster	Tennis	Golf	Baseball	Basketball	Football	Hockey	Total	Entropy	Purity
#1	1	1	0	11	4	676	693		
#2	27	89	333	827	253	33	1562		
#3	326	465	8	105	16	29	949		
Total	354	555	341	943	273	738	3204		

Select the collect entropy set of the confusion matrix.

- Cluster #1: 1.84, Cluster #2: 0.2, Cluster #3: 1.7, Total: 1.44
- Cluster #1: 1.44, Cluster #2: 1.84, Cluster #3: 1.7, Total: 0.2
- Cluster #1: 0.2, Cluster #2: 1.7, Cluster #3: 1.84, Total: 1.44
- Cluster #1: 0.2, Cluster #2: 1.44, Cluster #3: 1.7, Total: 1.84
- None of above

Cluster #1: 0.2, Cluster #2: 1.84, Cluster #3: 1.7, Total: 1.44

**Question 12****2 / 2 pts**

Review the table below.

Cluster	Tennis	Golf	Baseball	Basketball	Football	Hockey	Total	Entropy	Purity
#1	1	1	0	11	4	676	693		
#2	27	89	333	827	253	33	1562		
#3	326	465	8	105	16	29	949		
Total	354	555	341	943	273	738	3204		

Select the collect purity set of the confusion matrix

Cluster #1: 0.98, Cluster #2: 0.53, Cluster #3: 0.49, Total: 0.61

Cluster #1: 0.53, Cluster #2: 0.98, Cluster #3: 0.61, Total: 0.49

Cluster #1: 0.98, Cluster #2: 0.49, Cluster #3: 0.53, Total: 0.61

Cluster #1: 0.53, Cluster #2: 0.49, Cluster #3: 0.61, Total: 0.98

None of above

**Question 13****2 / 2 pts**

Review the following table.

	A	B	C	D	E
A	0	9	3	6	11
B	9	0	7	5	10
C	3	7	0	9	2
D	6	5	9	0	8
E	11	10	2	8	0

Select the correct DBSCAN clustering result with the parameters below

epsilon = 5, Minpts = 2

- cluster1: {A, C, E}, cluster2: {B,D}
- cluster1: {A}, cluster2: {B}, cluster3: {D}, and cluster4: {C, E}
- cluster1: {A, B}, cluster2: {C}, and cluster3: {D}, cluster4: {E}
- cluster1: {A}, cluster2: {B, E}, cluster3: {C}, and cluster4: {D}
- None of above

### Question 14

2 / 2 pts

Select qualities of clusters produced by a good clustering algorithm.

- Intra-cluster distances are minimized
- Inter-cluster distances are maximized
- Number of clusters produced
- A and B are correct

good cluster have strong cohesion within the cluster and maximum distance between them.

### Question 15

2 / 2 pts

What type of clustering is DBSCAN?

Center-based Contiguous Well-separated None of above

DBSCAN is a density-based clustering algorithm

**Question 16****2 / 2 pts**

Review the following table.

Data	X	Y
x1	1	1
x2	6	6
x3	7	7
x4	3	2
x5	6	8
x6	2	1
x7	2	3

Compute the average silhouette score to evaluate K-mean clustering result with the parameters below.

K:2

Initial centroids: centroid1 = (4,3) and centroid2 = (9,7)

Distance measure: Euclidean distance

0.752

**Question 17****2 / 2 pts**

Select all that are not true about DBSCAN.

- it has trouble when the clusters have widely varying densities
- it is very sensitive to noise and to the size of the data
- DBSCAN has a complexity of order  $O(n^3)$
- it can be expensive when the computation of nearest neighbors requires computing all pairwise proximities
- A and B are correct
- B and C are correct
  - DBSCAN is relatively resistant to noise and can handle clusters of arbitrary shapes and sizes. It has complexity order  $O(n \log n)$
- A, C and D are correct

**Question 18****2 / 2 pts**

Select the true statement for Nearest Neighbor classification



Although Nearest Neighbor classification has the data dimensionality issue, it can be solved by the scaling technique



k-NN classifier is a typical lazy learner because it spend so much time building the model even if the 'k' is very small



Drawing a circle is a mandatory task to find neighbor classes regardless of the number of 'K'



Determining the optimal value of 'K' is important for k-NN classifier performance.



None of above

### Question 19

2 / 2 pts

Review the table below.

Instance	X-Axis	Class
i	-3	-
ii	-2	-
iii	-1	+
iv	0	+
v	1	+
vi	2	-

Select all the support vector instances using Non-linear SVM

## Feature Map ( $\{R^1 \rightarrow R^2\}$ )

 i ii iii iv v vi

### Question 20

2 / 2 pts

Review the table below.

Instance	X-Axis	Class
i	-3	-
ii	-2	-
iii	-1	+
iv	0	+
v	1	+
vi	2	-

Select a correct hyper-plane equation of the data below using Non-linear SVM.

## Feature Map ( $\{R^1 \rightarrow R^2\}$ )

- Y=1
- Y=1.5
- Y=2
- Y=2.5
- None of above

Quiz Score: **40** out of 40

# Final Exam

**Due** Dec 9 at 1:15am      **Points** 40      **Questions** 20

**Available** Dec 6 at 12am - Dec 9 at 2am 3 days

**Time Limit** 75 Minutes

## Instructions

No collaboration

75 mins

## Attempt History

	<b>Attempt</b>	<b>Time</b>	<b>Score</b>
LATEST	<a href="#"><u>Attempt 1</u></a>	44 minutes	40 out of 40

! Correct answers are hidden.

Score for this quiz: **40** out of 40

Submitted Dec 8 at 6pm

This attempt took 44 minutes.

### Question 1

2 / 2 pts

Transaction ID Items Bought

- 1 {Milk, Beer, Diapers}
- 2 {Bread, Butter, Milk}
- 3 {Milk, Diapers, Cookies}
- 4 {Bread, Butter, Cookies}

- 5        {Beer, Cookies, Diapers}
- 6        {Milk, Diapers, Bread, Butter}
- 7        {Bread, Butter, Diapers}
- 8        {Beer, Diapers}
- 9        {Milk, Diapers, Bread, Butter}
- 10      {Beer, Cookies}

What is the maximum size of frequent itemsets that can be extracted  
(assuming minsup > 0.1)?

---

10

---

9

---

4

---

3

## Question 2

2 / 2 pts

Which of the above statements are true for any A, B, and C?

---

If A → B then B → A.

---

If A → B and B → C then A → C.

---

If A → C then A union B → C.

---

If A union B  $\rightarrow$  C then A  $\rightarrow$  C.

I & II

I, II, & III

I, II, & IV

II & III

none

### Question 3

2 / 2 pts

Example of market basket transactions.

#### Transaction ID Items Bought

1 {a, b, d, e}

2 {b, c, d}

3 {a, b, d, e}

4 {a, c, d, e}

5 {b, c, d, e}

6 {b, d, e}

7 {c, d}

8 {a, b, c}

9      {a, d, e}

10     {b, d}

Compute the confidence for the association rules  $\{a, e\} \rightarrow \{b\}$ .

0.5

**Question 4****2 / 2 pts**

Transaction ID Items Bought

1      {Milk, Beer, Diapers}

2      {Bread, Butter, Milk}

3      {Milk, Diapers, Cookies}

4      {Bread, Butter, Cookies}

5      {Beer, Cookies, Diapers}

6      {Milk, Diapers, Bread, Butter}

7      {Bread, Butter, Diapers}

8      {Beer, Diapers}

9      {Milk, Diapers, Bread, Butter}

10     {Beer, Cookies}

What is the maximum number of size 3-itemsets that can be derived from this data set. Note that the itemsets derived may not be in the dataset.

20

35

40

15

None of the above.

### Question 5

2 / 2 pts

Transaction ID Items Bought

- |   |                                |
|---|--------------------------------|
| 1 | {Milk, Beer, Diapers}          |
| 2 | {Bread, Butter, Milk}          |
| 3 | {Milk, Diapers, Cookies}       |
| 4 | {Bread, Butter, Cookies}       |
| 5 | {Beer, Cookies, Diapers}       |
| 6 | {Milk, Diapers, Bread, Butter} |
| 7 | {Bread, Butter, Diapers}       |
| 8 | {Beer, Diapers}                |

9 {Milk, Diapers, Bread, Butter}

10 {Beer, Cookies}

What is the confidence of the rule Beer -> Cookies

0.5

### Question 6

2 / 2 pts

Increase in size of a convolutional hidden layer would not necessarily increase the performance of a convolutional network.

True

False

### Question 7

2 / 2 pts

Consider a movie recommendation application, where an user can log in and select movies they like based on their previous preferences.

Which neural network architecture would be suitable to complete this task

Fully-Connected Neural Network.

Convolutional Neural Network.

Recurrent Neural Network.

Restricted Boltzmann Machine.

**Question 8****2 / 2 pts**

What is an autoencoder?

- a neural network that copies its input to its output.
- a neural network that codes itself.
- 
- a neural network that maps an output to an input through a hidden layer.
- a neural network that is trained to attempt to copy its input to its output.

**Question 9****2 / 2 pts**

Which of the following is an application of autoencoders?

- feature learning.
- dimensionality reduction.
- Rule mining
- all of the above.

**Question 10****2 / 2 pts**

Is an autoencoder the same as performing Principal Component Analysis?

- Not at all
- No but conceptually they can be used for the same purpose
- Autoencoders can do dimensionality reduction, but it is non-linear.
- Yes they have the same purpose.

**Question 11****2 / 2 pts**

Review the table below.

Cluster	Tennis	Golf	Baseball	Basketball	Football	Hockey	Total	Entropy	Purity
#1	1	1	0	11	4	676	693		
#2	27	89	333	827	253	33	1562		
#3	326	465	8	105	16	29	949		
Total	354	555	341	943	273	738	3204		

Select the collect entropy set of the confusion matrix.

- Cluster #1: 1.84, Cluster #2: 0.2, Cluster #3: 1.7, Total: 1.44
- Cluster #1: 1.44, Cluster #2: 1.84, Cluster #3: 1.7, Total: 0.2
- Cluster #1: 0.2, Cluster #2: 1.7, Cluster #3: 1.84, Total: 1.44
- Cluster #1: 0.2, Cluster #2: 1.44, Cluster #3: 1.7, Total: 1.84
- None of above

Cluster #1: 0.2, Cluster #2: 1.84, Cluster #3: 1.7, Total: 1.44

**Question 12****2 / 2 pts**

Review the table below.

Cluster	Tennis	Golf	Baseball	Basketball	Football	Hockey	Total	Entropy	Purity
#1	1	1	0	11	4	676	693		
#2	27	89	333	827	253	33	1562		
#3	326	465	8	105	16	29	949		
Total	354	555	341	943	273	738	3204		

Select the collect purity set of the confusion matrix

Cluster #1: 0.98, Cluster #2: 0.53, Cluster #3: 0.49, Total: 0.61

Cluster #1: 0.53, Cluster #2: 0.98, Cluster #3: 0.61, Total: 0.49

Cluster #1: 0.98, Cluster #2: 0.49, Cluster #3: 0.53, Total: 0.61

Cluster #1: 0.53, Cluster #2: 0.49, Cluster #3: 0.61, Total: 0.98

None of above

**Question 13****2 / 2 pts**

Review the following table.

	A	B	C	D	E
A	0	9	3	6	11
B	9	0	7	5	10
C	3	7	0	9	2
D	6	5	9	0	8
E	11	10	2	8	0

Select the correct DBSCAN clustering result with the parameters below

epsilon = 5, Minpts = 2

- cluster1: {A, C, E}, cluster2: {B,D}
- cluster1: {A}, cluster2: {B}, cluster3: {D}, and cluster4: {C, E}
- cluster1: {A, B}, cluster2: {C}, and cluster3: {D}, cluster4: {E}
- cluster1: {A}, cluster2: {B, E}, cluster3: {C}, and cluster4: {D}
- None of above

### Question 14

2 / 2 pts

Select qualities of clusters produced by a good clustering algorithm.

- Intra-cluster distances are minimized
- Inter-cluster distances are maximized
- Number of clusters produced
- A and B are correct

good cluster have strong cohesion within the cluster and maximum distance between them.

### Question 15

2 / 2 pts

What type of clustering is DBSCAN?

Center-based Contiguous Well-separated None of above

DBSCAN is a density-based clustering algorithm

**Question 16****2 / 2 pts**

Review the following table.

Data	X	Y
x1	1	1
x2	6	6
x3	7	7
x4	3	2
x5	6	8
x6	2	1
x7	2	3

Compute the average silhouette score to evaluate K-mean clustering result with the parameters below.

K:2

Initial centroids: centroid1 = (4,3) and centroid2 = (9,7)

Distance measure: Euclidean distance

0.752

**Question 17****2 / 2 pts**

Select all that are not true about DBSCAN.

- it has trouble when the clusters have widely varying densities
- it is very sensitive to noise and to the size of the data
- DBSCAN has a complexity of order  $O(n^3)$
- it can be expensive when the computation of nearest neighbors requires computing all pairwise proximities
- A and B are correct
- B and C are correct
  - DBSCAN is relatively resistant to noise and can handle clusters of arbitrary shapes and sizes. It has complexity order  $O(n \log n)$
- A, C and D are correct

**Question 18****2 / 2 pts**

Select the true statement for Nearest Neighbor classification



Although Nearest Neighbor classification has the data dimensionality issue, it can be solved by the scaling technique



k-NN classifier is a typical lazy learner because it spend so much time building the model even if the 'k' is very small



Drawing a circle is a mandatory task to find neighbor classes regardless of the number of 'K'



Determining the optimal value of 'K' is important for k-NN classifier performance.



None of above

### Question 19

2 / 2 pts

Review the table below.

Instance	X-Axis	Class
i	-3	-
ii	-2	-
iii	-1	+
iv	0	+
v	1	+
vi	2	-

Select all the support vector instances using Non-linear SVM

## Feature Map ( $\{R^1 \rightarrow R^2\}$ )

 i ii iii iv v vi

### Question 20

2 / 2 pts

Review the table below.

Instance	X-Axis	Class
i	-3	-
ii	-2	-
iii	-1	+
iv	0	+
v	1	+
vi	2	-

Select a correct hyper-plane equation of the data below using Non-linear SVM.

## Feature Map ( $\{R^1 \rightarrow R^2\}$ )

- Y=1
- Y=1.5
- Y=2
- Y=2.5
- None of above

Quiz Score: **40** out of 40

# Final Exam

**Due Dec 9 at 1:15am      Points 40      Questions 20**

**Available Dec 6 at 12am - Dec 9 at 2am 3 days**

**Time Limit 75 Minutes**

## Instructions

No collaboration

75 mins

## Attempt History

	<b>Attempt</b>	<b>Time</b>	<b>Score</b>
<b>LATEST</b>	<a href="#"><u>Attempt 1</u></a>	54 minutes	40 out of 40

! Correct answers are hidden.

Score for this quiz: **40** out of 40

Submitted Dec 8 at 5:57pm

This attempt took 54 minutes.

### Question 1

2 / 2 pts

Apriori pruning based on support is a greedy strategy but is not optimal

True

False

### Question 2

2 / 2 pts

Example of market basket transactions.

Transaction ID Items Bought

- |    |              |
|----|--------------|
| 1  | {a, b, d, e} |
| 2  | {b, c, d}    |
| 3  | {a, b, d, e} |
| 4  | {a, c, d, e} |
| 5  | {b, c, d, e} |
| 6  | {b, d, e}    |
| 7  | {c, d}       |
| 8  | {a, b, c}    |
| 9  | {a, d, e}    |
| 10 | {b, d}       |

Compute the support for itemset {b,c,d}.

0.2

**Question 3**

**2 / 2 pts**

Example of market basket transactions.

**Transaction ID Items Bought**

1 {a, b, d, e}

2 {b, c, d}

3 {a, b, d, e}

4 {a, c, d, e}

5 {b, c, d, e}

6 {b, d, e}

7 {c, d}

8 {a, b, c}

9 {a, d, e}

10 {b, d}

If the minimum support is set at 40%, how many frequent 3-itemsets will be found? Note that if a 3 itemset is a subset of a larger itemset, it counts as one occurrence.

2

**Question 4**

2 / 2 pts

Select the reason why Apriori pruning in the search for frequent itemsets works.

- support count is monotonic with respect to itemsets.
- we search in transaction ID order.
- support count diverges as we add to the itemset.
- support count is anti-monotonic with respect to itemsets.
- it is an excellent heuristic, but it does not work 100% of the time.

**Question 5****2 / 2 pts****Transaction ID Items Bought**

- 1 {Milk, Beer, Diapers}
- 2 {Bread, Butter, Milk}
- 3 {Milk, Diapers, Cookies}
- 4 {Bread, Butter, Cookies}
- 5 {Beer, Cookies, Diapers}
- 6 {Milk, Diapers, Bread, Butter}
- 7 {Bread, Butter, Diapers}
- 8 {Beer, Diapers}
- 9 {Milk, Diapers, Bread, Butter}
- 10 {Beer, Cookies}

Find all itemsets (of size 3 or larger) that has the largest support.

{Milk, Beer, Diapers}

{Bread, Butter, Milk}

{Bread, Butter, Diapers}

{Milk, Diapers, Bread}

None of the above.

All of the above.

## Question 6

2 / 2 pts

Usage of deep learning is predicated with availability of large amounts of data. Is the data volume the only determining factor to check whether we can use deep learning ?

Yes, lower volumes lead to overfitting

No, balance among classes in the input data is the only determining factor  
volume does not matter

No, balance among classes in the input data is a co-factor along with the  
data volume

None of the above

**Question 7****2 / 2 pts**

Consider a movie recommendation application, where an user can log in and select movies they like based on their previous preferences.

Which neural network architecture would be suitable to complete this task

- 
- Fully-Connected Neural Network.
  - Convolutional Neural Network.
  - Recurrent Neural Network.
  - Restricted Boltzmann Machine.
- 

**Question 8****2 / 2 pts**

Which of the following is an application of autoencoders?

- 
- feature learning.
  - dimensionality reduction.
  - Rule mining
  - all of the above.
- 

**Question 9****2 / 2 pts**

What is an autoencoder?

- a neural network that copies its input to its output.
- a neural network that codes itself.
- 
- a neural network that maps an output to an input through a hidden layer.
- a neural network that is trained to attempt to copy its input to its output.

**Question 10****2 / 2 pts**

Increase in size of a convolutional hidden layer would not necessarily increase the performance of a convolutional network.

- True
- False

**Question 11****2 / 2 pts**

Review the table below.

Cluster	Tennis	Golf	Baseball	Basketball	Football	Hockey	Total	Entropy	Purity
#1	1	1	0	11	4	676	693		
#2	27	89	333	827	253	33	1562		
#3	326	465	8	105	16	29	949		
Total	354	555	341	943	273	738	3204		

Select the collect entropy set of the confusion matrix.

- Cluster #1: 1.84, Cluster #2: 0.2, Cluster #3: 1.7, Total: 1.44

Cluster #1: 1.44, Cluster #2: 1.84, Cluster #3: 1.7, Total: 0.2

Cluster #1: 0.2, Cluster #2: 1.7, Cluster #3: 1.84, Total: 1.44

Cluster #1: 0.2, Cluster #2: 1.44, Cluster #3: 1.7, Total: 1.84

None of above

Cluster #1: 0.2, Cluster #2: 1.84, Cluster #3: 1.7, Total: 1.44

## Question 12

2 / 2 pts

Select qualities of clusters produced by a good clustering algorithm.

Intra-cluster distances are minimized

Inter-cluster distances are maximized

Number of clusters produced

A and B are correct

good cluster have strong cohesion within the cluster and maximum distance between them.

## Question 13

2 / 2 pts

Select all statements that are true for Hierarchical clustering.

Although it requires high storage, high computing is not required

Robust to noise Robust to outliers

It is not easy to decide on the number of cluster because the complicated dendrogram disturbs this decision

 None of above

### Question 14

2 / 2 pts

Review the following table.

Data	X	Y
A	1	2
B	2	2
C	2	3
D	8	7
E	8	8
F	8	2

Select the correct number of core points and DBSCAN clustering result with the parameters below.

Eps: 5

Minimum samples (MinPts): 3

Distance measurement: Euclidean distance

 5, cluster1: {A,B,C}, cluster2: {D,E,F} 6, cluster1: {A,B}, cluster2: {C,D,E}, cluster3: {F} 4, cluster1: {A,B,C}, cluster2: {D, E, F}

- 5, cluster1: {A}, cluster2: {B}, cluster3: {C}, cluster4: {D}, cluster5: {E}
- None of above

**Question 15****2 / 2 pts**

Review the following table.

Data	X	Y
x1	1	1
x2	6	6
x3	7	7
x4	3	2
x5	6	8
x6	2	1
x7	2	3

Compute the sum of squared error (SSE) to evaluate K-mean clustering result with the parameters below.

K:2

Initial centroids: centroid1 = (4,3) and centroid2 = (9,7)

Distance measure: Euclidean distance

7.4167

**Question 16****2 / 2 pts**

Select all that are not true about DBSCAN.

- it has trouble when the clusters have widely varying densities
- it is very sensitive to noise and to the size of the data
- DBSCAN has a complexity of order  $O(n^3)$
- it can be expensive when the computation of nearest neighbors requires computing all pairwise proximities
- A and B are correct
- B and C are correct
  - DBSCAN is relatively resistant to noise and can handle clusters of arbitrary shapes and sizes. It has complexity order  $O(n \log n)$
- A, C and D are correct

**Question 17****2 / 2 pts**

Select all different types of clusters.

- Well-separated
- Center-based
- Contiguous
- Density-based
- All of the above

all are type of clustering

**Question 18**

2 / 2 pts

Review the table below.

Instance	X-Axis	Class
i	-3	-
ii	-2	-
iii	-1	+
iv	0	+
v	1	+
vi	2	-

Select all the support vector instances using Non-linear SVM

Feature Map ( $\{R^1 \rightarrow R^2\}$ )

---

 i

---

 ii

---

 iii

---

 iv

---

 v

---

 vi

**Question 19****2 / 2 pts**

Review the table below.

A	B	C	Class
F	F	F	Y
F	F	T	N
F	T	T	N
F	T	T	N
F	F	T	Y
T	F	T	Y
T	F	T	N
T	F	T	N
T	T	T	Y
T	F	T	Y

Select the correct prediction class for a test sample (A=F, B=T, C=F) using the naïve Bayes approach.

---

 'Y'

---

 'N'

---

 Anyone**Question 20****2 / 2 pts**

Review the table below.

Rainy	Sunshine		Wind		Atmosphere		Temperature	
	Yes	No	Yes	No	High	Low	High	Low
Yes	2	6	6	2	8	0	5	3
No	8	4	2	10	2	10	6	6

Is it raining if the weather condition is *sunshine, no wind, high atmosphere*, and *low temperature* (*Hint: you should use Naïve Bayes*)?

Yes

No

Quiz Score: **40** out of 40

# Final Exam

**Due** Dec 8 at 1:15am      **Points** 40      **Questions** 20

**Available** Dec 6 at 12am - Dec 8 at 2am 2 days

**Time Limit** 75 Minutes

## Instructions

No collaboration

75 mins

## Attempt History

	<b>Attempt</b>	<b>Time</b>	<b>Score</b>
<b>LATEST</b>	<a href="#"><u>Attempt 1</u></a>	50 minutes	38 out of 40

! Correct answers are hidden.

Score for this quiz: **38** out of 40

Submitted Dec 6 at 2:37pm

This attempt took 50 minutes.

<b>Question 1</b>		<b>2 / 2 pts</b>
Example of market basket transactions.		
Transaction ID Items Bought		
1	{a, b, d, e}	
2	{b, c, d}	
3	{a, b, d, e}	

- 4      {a, c, d, e}
- 5      {b, c, d, e}
- 6      {b, d, e}
- 7      {c, d}
- 8      {a, b, c}
- 9      {a, d, e}
- 10     {b, d}

Compute the support for itemset {b,c,d}.

0.2

## Question 2

2 / 2 pts

Example of market basket transactions.

### Transaction ID Items Bought

- 1      {a, b, d, e}
- 2      {b, c, d}
- 3      {a, b, d, e}
- 4      {a, c, d, e}

5      {b, c, d, e}

6      {b, d, e}

7      {c, d}

8      {a, b, c}

9      {a, d, e}

10     {b, d}

Compute the confidence for the association rules  $\{a, e\} \rightarrow \{b\}$ .

0.5

### Question 3

2 / 2 pts

Transaction ID Items Bought

1      {Milk, Beer, Diapers}

2      {Bread, Butter, Milk}

3      {Milk, Diapers, Cookies}

4      {Bread, Butter, Cookies}

5      {Beer, Cookies, Diapers}

6      {Milk, Diapers, Bread, Butter}

- 7 {Bread, Butter, Diapers}
- 8 {Beer, Diapers}
- 9 {Milk, Diapers, Bread, Butter}
- 10 {Beer, Cookies}

What is the confidence of the rule Beer  $\rightarrow$  Cookies

0.5

#### Question 4

2 / 2 pts

Example of market basket transactions.

#### Transaction ID Items Bought

- 1 {a, b, d, e}
- 2 {b, c, d}
- 3 {a, b, d, e}
- 4 {a, c, d, e}
- 5 {b, c, d, e}
- 6 {b, d, e}
- 7 {c, d}

8            {a, b, c}

9            {a, d, e}

10          {b, d}

If the minimum support is set at 40%, how many frequent 3-itemsets will be found? Note that if a 3 itemset is a subset of a larger itemset, it counts as one occurrence.

2

### Question 5

2 / 2 pts

Transaction ID Items Bought

1            {Milk, Beer, Diapers}

2            {Bread, Butter, Milk}

3            {Milk, Diapers, Cookies}

4            {Bread, Butter, Cookies}

5            {Beer, Cookies, Diapers}

6            {Milk, Diapers, Bread, Butter}

7            {Bread, Butter, Diapers}

8            {Beer, Diapers}

9 {Milk, Diapers, Bread, Butter}

10 {Beer, Cookies}

What is the maximum number of size 3-itemsets that can be derived from this data set. Note that the itemsets derived may not be in the dataset.

20

35

40

15

None of the above.

### Question 6

2 / 2 pts

Which of the following is an application of autoencoders?

feature learning.

dimensionality reduction.

Rule mining

all of the above.

### Question 7

2 / 2 pts

The number of neurons in the output layer should match the number of classes (Where the number of classes is greater than 2) in a supervised learning task.

---

True

---

False

**Question 8****2 / 2 pts**

What is an autoencoder?

---

a neural network that copies its input to its output.

---

a neural network that codes itself.

---

a neural network that maps an output to an input through a hidden layer.

---

a neural network that is trained to attempt to copy its input to its output.

**Question 9****2 / 2 pts**

Consider a movie recommendation application, where an user can log in and select movies they like based on their previous preferences.

Which neural network architecture would be suitable to complete this task

---

Fully-Connected Neural Network.

---

Convolutional Neural Network.

---

- Recurrent Neural Network.
- 
- Restricted Boltzmann Machine.

Incorrect

**Question 10****0 / 2 pts**

Increase in size of a convolutional hidden layer would not necessarily increase the performance of a convolutional network.

- 
- True
- 
- False

**Question 11****2 / 2 pts**

Select different aspects of cluster validation.

- 
- Determining the clustering tendency of a set of data
- 
- To improve the complexity of the algorithm
- 
- Determining the 'correct' number of clusters
- 
- None of the above

correct these are some of the aspects of cluster validation!

**Question 12****2 / 2 pts**

What is the purpose of cluster analysis?

- To avoid finding patterns in noise
- To compare clustering algorithms
- To compare two sets of clusters
- all of the above
- None of the above

we do cluster validity to avoid clustering noise and find a suitable algorithm for our data

**Question 13****2 / 2 pts**

Select qualities of clusters produced by a good clustering algorithm.

- Intra-cluster distances are minimized
- Inter-cluster distances are maximized
- Number of clusters produced
- A and B are correct

good cluster have strong cohesion within the cluster and maximum distance between them.

**Question 14****2 / 2 pts**

Select all statements that are true for Hierarchical clustering.

Although it requires high storage, high computing is not required

Robust to noise

Robust to outliers

It is not easy to decide on the number of cluster because the complicated dendrogram disturbs this decision

None of above

**Question 15****2 / 2 pts**

Review the following table.

Data	X	Y
A	1	2
B	2	2
C	2	3
D	8	7
E	8	8
F	8	2

Select the correct number of core points and DBSCAN clustering result with the parameters below.

Eps: 5

Minimum samples (MinPts):3

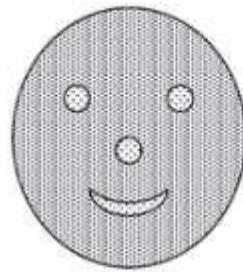
### Distance measurement: Euclidean distance

- 5, cluster1: {A,B,C}, cluster2: {D,E,F}
- 6, cluster1: {A,B}, cluster2: {C,D,E}, cluster3: {F}
- 4, cluster1: {A,B,C}, cluster2: {D, E, F}
- 5, cluster1: {A}, cluster2: {B}, cluster3: {C}, cluster4: {D}, cluster5: {E}
- None of above

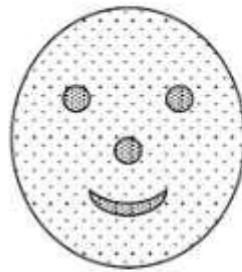
### Question 16

2 / 2 pts

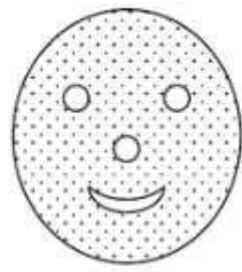
Review the image below.



(a)



(b)



(c)



(d)

**Darkness or number of dots represent density. Lines are used only to distinguish regions and do not represent points.**

If you want to find the patterns represented by the nose, eyes, and mouth using k-Means clustering, select all figures that be well-clustered?

 (a) (b) (c)

(d) None of above**Question 17****2 / 2 pts**

Select all that are not true about DBSCAN.

- it has trouble when the clusters have widely varying densities
- it is very sensitive to noise and to the size of the data
- DBSCAN has a complexity of order  $O(n^3)$
- 
- it can be expensive when the computation of nearest neighbors requires computing all pairwise proximities
- A and B are correct
- B and C are correct

DBSCAN is relatively resistant to noise and can handle clusters of arbitrary shapes and sizes. It has complexity order  $O(n \log n)$

- A, C and D are correct

**Question 18****2 / 2 pts**

Review the table below.

Instance	X-Axis	Class
i	-3	-
ii	-2	-
iii	-1	+
iv	0	+
v	1	+
vi	2	-

Select a correct hyper-plane equation of the data below using Non-linear SVM.

## Feature Map ( $\{R^1 \rightarrow R^2\}$ )

Y=1

Y=1.5

Y=2

Y=2.5

None of above

### Question 19

2 / 2 pts

Review the table below.

Instance	X-Axis	Class
i	-3	-
ii	-2	-
iii	-1	+
iv	0	+
v	1	+
vi	2	-

Select all the support vector instances using Non-linear SVM

## Feature Map ( $\{R^1 \rightarrow R^2\}$ )

i

ii

iii

iv

v

vi

### Question 20

2 / 2 pts

Review the table below.

A	B	C	Class
F	F	F	Y
F	F	T	N
F	T	T	N
F	T	T	N
F	F	T	Y
T	F	T	Y
T	F	T	N
T	F	T	N
T	T	T	Y
T	F	T	Y

Select the correct prediction class for a test sample (A=F, B=T, C=F) using the naïve Bayes approach.

'Y'

'N'

Anyone

Quiz Score: **38** out of 40

# Final Exam

**Due** Dec 9 at 1:15am      **Points** 40      **Questions** 20

**Available** Dec 6 at 12am - Dec 9 at 2am 3 days

**Time Limit** 75 Minutes

## Instructions

No collaboration

75 mins

## Attempt History

	Attempt	Time	Score
LATEST	<a href="#"><u>Attempt 1</u></a>	39 minutes	40 out of 40

❗ Correct answers are hidden.

Score for this quiz: **40** out of 40

Submitted Dec 8 at 5:10pm

This attempt took 39 minutes.

Question 1	2 / 2 pts
Select the reason why Apriori pruning in the search for frequent itemsets works.	
<input type="radio"/> support count is monotonic with respect to itemsets.	
<input type="radio"/> we search in transaction ID order.	
<input type="radio"/> support count diverges as we add to the itemset.	
<input checked="" type="radio"/> support count is anti-monotonic with respect to itemsets.	

- it is an excellent heuristic, but it does not work 100% of the time.

**Question 2****2 / 2 pts**

Example of market basket transactions.

Transaction ID Items Bought

1 {a, b, d, e}

2 {b, c, d}

3 {a, b, d, e}

4 {a, c, d, e}

5 {b, c, d, e}

6 {b, d, e}

7 {c, d}

8 {a, b, c}

9 {a, d, e}

10 {b, d}

Compute the support for itemset {b,c,d}.

0.2

**Question 3**

**2 / 2 pts**

## Transaction ID Items Bought

- 1 {Milk, Beer, Diapers}
- 2 {Bread, Butter, Milk}
- 3 {Milk, Diapers, Cookies}
- 4 {Bread, Butter, Cookies}
- 5 {Beer, Cookies, Diapers}
- 6 {Milk, Diapers, Bread, Butter}
- 7 {Bread, Butter, Diapers}
- 8 {Beer, Diapers}
- 9 {Milk, Diapers, Bread, Butter}
- 10 {Beer, Cookies}

What is the confidence of the rule Beer  $\rightarrow$  Cookies

0.5

**Question 4****2 / 2 pts**

Example of market basket transactions.

## Transaction ID Items Bought

1 {a, b, d, e}

2 {b, c, d}

3 {a, b, d, e}

4 {a, c, d, e}

5 {b, c, d, e}

6 {b, d, e}

7 {c, d}

8 {a, b, c}

9 {a, d, e}

10 {b, d}

If the minimum support is set at 40%, how many frequent 3-itemsets will be found? Note that if a 3 itemset is a subset of a larger itemset, it counts as one occurrence.

2

**Question 5****2 / 2 pts**

Apriori pruning based on support is a greedy strategy but is not optimal

True False**Question 6****2 / 2 pts**

Usage of deep learning is predicated with availability of large amounts of data. Is the data volume the only determining factor to check whether we can use deep learning ?

 Yes, lower volumes lead to overfitting

No, balance among classes in the input data is the only determining factor volume does not matter

No, balance among classes in the input data is a co-factor along with the data volume

 None of the above**Question 7****2 / 2 pts**

The number of neurons in the output layer should match the number of classes (Where the number of classes is greater than 2) in a supervised learning task.

 True False

**Question 8****2 / 2 pts**

Is an autoencoder the same as performing Principal Component Analysis?

- 
- Not at all
- 
- No but conceptually they can be used for the same purpose
- 
- Autoencoders can do dimensionality reduction, but it is non-linear.
- 
- Yes they have the same purpose.

**Question 9****2 / 2 pts**

Consider a movie recommendation application, where an user can log in and select movies they like based on their previous preferences.

Which neural network architecture would be suitable to complete this task

- 
- Fully-Connected Neural Network.
- 
- Convolutional Neural Network.
- 
- Recurrent Neural Network.
- 
- Restricted Boltzmann Machine.

**Question 10****2 / 2 pts**

Increase in size of a convolutional hidden layer would not necessarily increase the performance of a convolutional network.

True

False

### Question 11

2 / 2 pts

Review the table below.

Cluster	Tennis	Golf	Baseball	Basketball	Football	Hockey	Total	Entropy	Purity
#1	1	1	0	11	4	676	693		
#2	27	89	333	827	253	33	1562		
#3	326	465	8	105	16	29	949		
Total	354	555	341	943	273	738	3204		

Select the collect purity set of the confusion matrix

Cluster #1: 0.98, Cluster #2: 0.53, Cluster #3: 0.49, Total: 0.61

Cluster #1: 0.53, Cluster #2: 0.98, Cluster #3: 0.61, Total: 0.49

Cluster #1: 0.98, Cluster #2: 0.49, Cluster #3: 0.53, Total: 0.61

Cluster #1: 0.53, Cluster #2: 0.49, Cluster #3: 0.61, Total: 0.98

None of above

### Question 12

2 / 2 pts

Review the following table.

Data	X	Y
x1	1	1
x2	6	6
x3	7	7
x4	3	2
x5	6	8
x6	2	1
x7	2	3

Select the correct pair of final centroid and clustering set using K-mean clustering with the parameters below.

K:2

Initial centroids: centroid1 = (4,3) and centroid2 = (9,7)

Distance measure: Euclidean distance



Cluster 1: (x1, x4, x6, x7) with centroid=(2, 1.75), Cluster 2: (x2, x3, x5) with centroid=(6.33, 7)



Cluster 1: (x1, x4, x6, x7) with centroid=(1.75, 2), Cluster 2: (x2, x3, x5) with centroid=(6.33, 7)



Cluster 1: (x1, x4, x6, x7) with centroid=(2, 1.75), Cluster 2: (x2, x3, x5) with centroid=(7, 6.33)



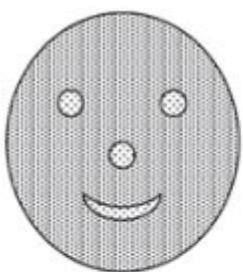
Cluster 1: (x1, x2, x4, x6, x7) with centroid=(2, 1.75), Cluster 2: (x3, x5) with centroid=(6.33, 7)



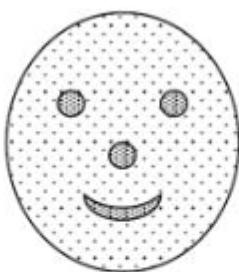
Cluster 1: (x1, x2, x4, x6, x7) with centroid=(2, 1.75), Cluster 2: (x3, x5) with centroid=(7, 6.33)

**Question 13****2 / 2 pts**

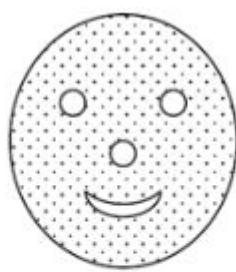
Review the image below.



(a)



(b)



(c)



(d)

**Darkness or number of dots represent density. Lines are used only to distinguish regions and do not represent points.**

If you want to find the patterns represented by the nose, eyes, and mouth using k-Means clustering, select all figures that be well-clustered?

 (a) (b) (c) (d) None of above**Question 14****2 / 2 pts**

Select qualities of clusters produced by a good clustering algorithm.

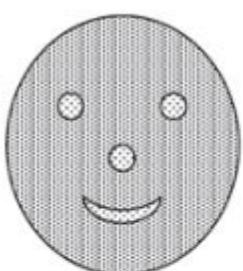
- Intra-cluster distances are minimized
- Inter-cluster distances are maximized
- Number of clusters produced
- A and B are correct

good cluster have strong cohesion within the cluster and maximum distance between them.

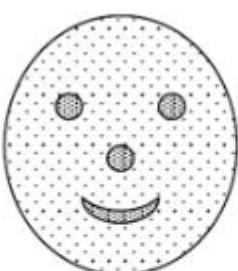
### Question 15

2 / 2 pts

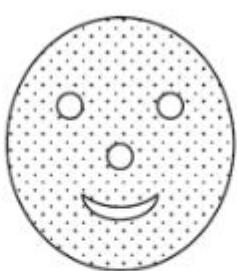
Review the image below.



(a)



(b)



(c)



(d)

**Darkness or number of dots represent density. Lines are used only to distinguish regions and do not represent points.**

If we want to find the patterns represented by the nose, eyes, and mouth using single linkage hierarchical clustering, select all figures that be well-clustered?

(a)

(b)

(c)

(d) None of above**Question 16****2 / 2 pts**

Select all different type of clusters.

 Well-separated Center-based Contiguous Density-based All of the above

all are type of clustering

**Question 17****2 / 2 pts**

Review the following table.

	A	B	C	D	E
A	0	9	3	6	11
B	9	0	7	5	10
C	3	7	0	9	2
D	6	5	9	0	8
E	11	10	2	8	0

Select the correct DBSCAN clustering result with the parameters below

epsilon = 5, Minpts = 2

- cluster1: {A, C, E}, cluster2: {B,D}
- cluster1: {A}, cluster2: {B}, cluster3: {D}, and cluster4: {C, E}
- cluster1: {A, B}, cluster2: {C}, and cluster3: {D}, cluster4: {E}
- cluster1: {A}, cluster2: {B, E}, cluster3: {C}, and cluster4: {D}
- None of above

### Question 18

2 / 2 pts

Review the table below.

A	B	C	Class
F	F	F	Y
F	F	T	N
F	T	T	N
F	T	T	N
F	F	T	Y
T	F	T	Y
T	F	T	N
T	F	T	N
T	T	T	Y
T	F	T	Y

Select the correct prediction class for a test sample (A=F, B=T, C=F) using the naïve Bayes approach.

'Y'

'N'

Anyone

### Question 19

2 / 2 pts

Select the true statement for Nearest Neighbor classification

Although Nearest Neighbor classification has the data dimensionality issue, it can be solved by the scaling technique

k-NN classifier is a typical lazy learner because it spend so much time building the model even if the 'k' is very small

Drawing a circle is a mandatory task to find neighbor classes regardless of the number of 'K'

Determining the optimal value of 'K' is important for k-NN classifier performance.

None of above

### Question 20

2 / 2 pts

Review the table below.

Instance	X-Axis	Class
i	-3	-
ii	-2	-
iii	-1	+
iv	0	+
v	1	+
vi	2	-

Select all the support vector instances using Non-linear SVM

Feature Map ( $\{R^1 \rightarrow R^2\}$ )

 i

<input checked="" type="checkbox"/> ii
<input checked="" type="checkbox"/> iii
<input type="checkbox"/> iv
<input checked="" type="checkbox"/> v
<input checked="" type="checkbox"/> vi

Quiz Score: **40** out of 40

# Midterm

---

**Due** Oct 28 at 11:59pm      **Points** 48

**Questions** 16

**Available** Oct 26 at 12am - Oct 29 at 1am 3  
days

**Time Limit** 100 Minutes

---

## Attempt History

	<b>Attempt</b>	<b>Time</b>	<b>Score</b>
<b>LATEST</b>	<a href="#"><u>Attempt 1</u></a>	40 minutes	45 out of 48

---

Score for this quiz: **45** out of 48

Submitted Oct 27 at 6:34pm

This attempt took 40 minutes.

### Question 1

3 / 3 pts

Select the correct sentence or sentences that define cosine and correlation measures.



The range of values that are possible for the cosine measure is [0, 1]



If two objects have a cosine similarity of 1, they are identical

Correct!



If X and Y have a mean of 0,  $\text{corr}(X,Y) = \cos(X,Y)$



A, B



A, B and C

## Question 2

3 / 3 pts

Select all possible use-cases of clustering



Credit card fraud detection



Market segmentation

orrect!

Summarize news

A, B, and C

A, B

### Question 3

3 / 3 pts

Select all predictive techniques.

orrect!

Anomaly Detection

orrect!

Regression

orrect!

Classification

Clustering

Dimensionality Reduction

### Question 4

3 / 3 pts

Select all unsupervised learning applications

orrect!



Marketing and sales promotion using Association Rule

Association Rule Mining is an unsupervised learning method.

orrect!



Document clustering

Clustering is an unsupervised learning method.



Face recognition



Fraud detection in credit card transactions

## Question 5

3 / 3 pts

Select all correct sentences that define correlation.

orrect!



Correlation -1 means that the vectors have a perfect negative linear relationship.

Correct. If the correlation is -1, the vectors follow the perfect negative linear correlation.



Correlation 1 means that the vectors have a perfect negative linear relationship.

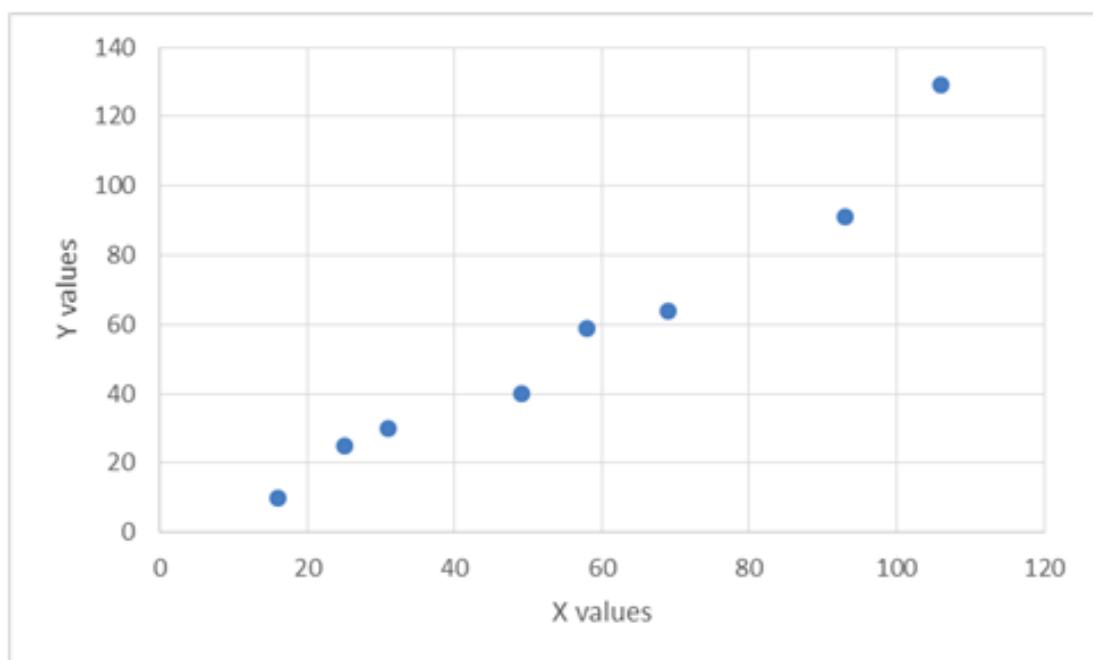


Correlation 0 means that the vectors have a perfect positive linear relationship.

## Question 6

3 / 3 pts

	X	Y
A	16	10
B	25	25
C	31	30
D	49	40
E	58	59
F	69	64
G	93	91
H	106	129



Compute cosine similarity between X and Y

orrect!

0.991

ect Answers

Between 0.95 and 1

## Question 7

3 / 3 pts

Select all techniques used to measure node impurity.

orrect!

Gini Index

Gini Index is a measure of node impurity

average

orrect!

Entropy

Entropy is a measure of node impurity

split

## Question 8

3 / 3 pts

Select all sentences that correctly define errors.

orrect!



Prediction errors consist of bias, variance, and irreducible error.

orrect!



Bias error results in high training error.



Variance error results in high training error.

orrect!



Irreducible error cannot be optimized.

## Question 9

3 / 3 pts

Review the table below.

ID	Device OS	Carrier	Device Color	Earlier Adopter (Class)
1	Android	T-Mobile	White	Y
2	Android	AT&T	Gray	Y
3	Android	AT&T	Gray	Y
4	Android	AT&T	Black	Y
5	Android	AT&T	Pink	Y
6	Android	AT&T	Pink	Y
7	IOS	AT&T	White	Y
8	IOS	AT&T	White	Y
9	IOS	AT&T	Gray	Y
10	IOS	Verizon	Black	Y
11	Android	T-Mobile	Black	N
12	Android	T-Mobile	Pink	N
13	Android	T-Mobile	Gray	N
14	Android	Verizon	Pink	N
15	IOS	Verizon	White	N
16	IOS	Verizon	White	N
17	IOS	Verizon	Gray	N
18	IOS	Verizon	Gray	N
19	IOS	Verizon	Gray	N
20	IOS	Verizon	Black	N

Based on the table, which of the following statements are correct about the decision tree model?

orrect!



The ID is the best attribute to predict Earlier Adopter



The Carrier is the best attribute to predict Earlier Adopter



The Device OS is the best attribute to predict Earlier Adopter



The Device Color is the best attribute to predict Earlier Adopter

## Question 10

3 / 3 pts

Select Methods for good model evaluation.

orrect!



Maximize values of both precision and recall

we need to maximize both precision and recall

Maximize accuracy

Maximize precision

Minimize recall

## Question 11

3 / 3 pts

Review the following table.

Ground Truth	apple									
Predict	apple	pear	pear	apple	apple	pear	pear	pear	apple	apple

Compute F1 for 'pear' class

orrect!

0

ect Answers

0 (with margin: 0)

## Question 12

3 / 3 pts

Select all sentences that are correct about cross-validation.

orrect!

Leave-one-out is a type of cross-validation.



Cross-validation guarantees reduction in test error.



K-fold cross-validation requires K repetitions of the same train set on K different test sets..

orrect!



In cross validation the same training record can be selected multiple times.

## Question 13

3 / 3 pts

Review the following table.

Ground Truth	apple									
Predict	apple	pear	pear	apple	apple	pear	pear	pear	apple	apple

Compute the accuracy as a fraction

orrect!

0.5

ect Answers

0.5 (with margin: 0)

### Question 14

3 / 3 pts

Review the following table.

Instance	Attribute 1	Attribute 2	Attribute 3	class
1	T	T	1	Y
2	T	T	6	Y
3	T	F	5	N
4	F	F	4	Y
5	F	T	7	N
6	F	T	3	N
7	F	F	8	N
8	T	F	7	Y
9	F	T	5	N

What is the best split between Attribute 1 and Attribute 2 according to the Gini index? What is the Gini index of the split attribute?

orrect!



Attribute 1 is the best split and Gini index is 0.3444



Attribute 1 is the best split and Gini index is 0.4889



Attribute 2 is the best split and Gini index is 0.3444



Attribute 2 is the best split and Gini index is 0.4889

### Question 15

3 / 3 pts

Review the following table.

Instance	Attribute 1	Attribute 2	Attribute 3	class
1	T	T	1	Y
2	T	T	6	Y
3	T	F	5	N
4	F	F	4	Y
5	F	T	7	N
6	F	T	3	N
7	F	F	8	N
8	T	F	7	Y
9	F	T	5	N

Select the best split among attribute 1, attribute 2, and attribute 3, according to the information gain.

orrect!

Attribute 1

Attribute 2

Attribute 3

## Question 16

0 / 3 pts

ID	Device OS	Carrier	Device Color	Earlier Adopter (Class)
1	Android	T-Mobile	White	Y
2	Android	AT&T	Gray	Y
3	Android	AT&T	Gray	Y
4	Android	AT&T	Black	Y
5	Android	AT&T	Pink	Y
6	Android	AT&T	Pink	Y
7	IOS	AT&T	White	Y
8	IOS	AT&T	White	Y
9	IOS	AT&T	Gray	Y
10	IOS	Verizon	Black	Y
11	Android	T-Mobile	Black	N
12	Android	T-Mobile	Pink	N
13	Android	T-Mobile	Gray	N
14	Android	Verizon	Pink	N
15	IOS	Verizon	White	N
16	IOS	Verizon	White	N
17	IOS	Verizon	Gray	N
18	IOS	Verizon	Gray	N
19	IOS	Verizon	Gray	N
20	IOS	Verizon	Black	N

Compute the Gini index for Device Color attribute using multiway split

Answered

0.492

Correct Answers

Between 0.49 and 0.4915

Quiz Score: **45** out of 48

# Midterm

**Due** Oct 28 at 11:59pm      **Points** 48      **Questions** 16

**Available** Oct 26 at 12am - Oct 29 at 1am 3 days

**Time Limit** 100 Minutes

## Attempt History

	<b>Attempt</b>	<b>Time</b>	<b>Score</b>
<b>LATEST</b>	<a href="#"><u>Attempt 1</u></a>	13 minutes	47 out of 48

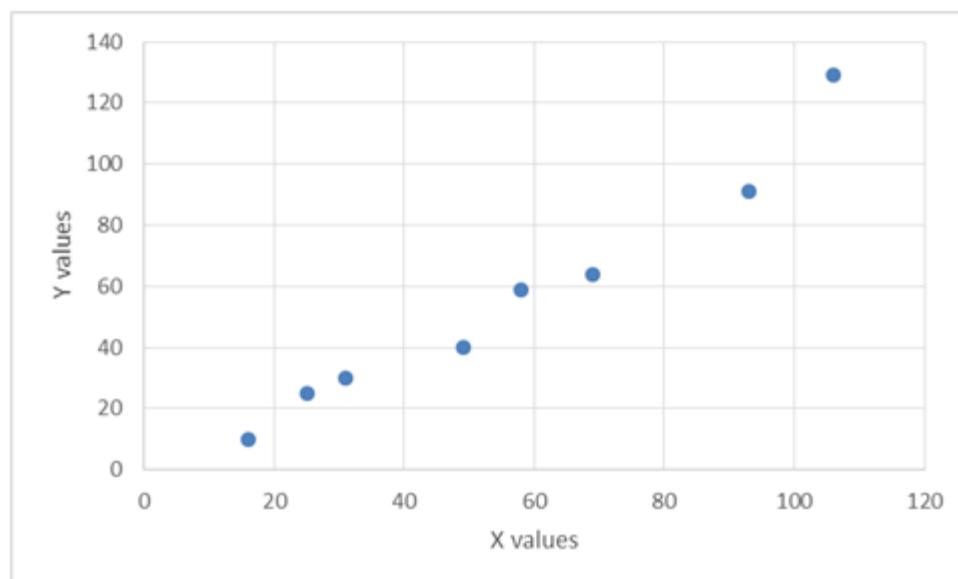
Score for this quiz: **47** out of 48

Submitted Oct 27 at 3:23pm

This attempt took 13 minutes.

<b>Question 1</b>			<b>3 / 3 pts</b>

	X	Y
A	16	10
B	25	25
C	31	30
D	49	40
E	58	59
F	69	64
G	93	91
H	106	129



Compute cosine similarity between X and Y

Correct!

0.973

Incorrect Answers

Between 0.95 and 1

## Question 2

3 / 3 pts

Select all correct sentences which make up the clustering.

- K-Nearest Neighbor is a typical clustering technique
- Association rule discovery is a typical technique of clustering
- Clustering techniques do not require any parameters
- A, B
- A, B, and C

Correct!

**Question 3****3 / 3 pts**

Select all possible use-cases of clustering

- Credit card fraud detection
- Market segmentation
- Summarize news

**Correct!**

- A, B, and C
- A, B

**Question 4****2 / 3 pts**

Select all predictive techniques.

**Incorrect Answer**

- Anomaly Detection

**Correct!**

- Regression

**Correct!**

- Classification

- Clustering

- Dimensionality Reduction

**Question 5****3 / 3 pts**

Select the correct sentence or sentences that define cosine and correlation

measures.

The range of values that are possible for the cosine measure is [0, 1]

If two objects have a cosine similarity of 1, they are identical

**Correct!**

If  $X$  and  $Y$  have a mean of 0,  $\text{corr}(X,Y) = \cos(X,Y)$

A, B

A, B and C

### Question 6

3 / 3 pts

Tid	Refund	Marital Status	Taxable Income	Cheat
1	Yes	Single	125K	No
2	No	Married	100K	No
3	No	Single	70K	No
4	Yes	Married	120K	No
5	No	Divorced	95K	Yes
6	No	Married	60K	No
7	Yes	Divorced	220K	No
8	No	Single	85K	Yes
9	No	Married	75K	No
10	No	Single	90K	Yes

Select all possible columns which can be set as the class columns.

Tid

**Correct!**

Refund

**Correct!**

Marital Status

**Correct!** Taxable Income**Correct!** Cheat**Question 7****3 / 3 pts**

Select the statements below that are computed as part of the performance metric.

**Correct!** Precision can be computed using True Positive and False Positive. Recall can be computed using True Positive and False Positive.**Correct!** F1 can be computed using True Positive and False Positive, and False Negative. Accuracy can be computed using True Positive, False Positive, and False Negative.**Question 8****3 / 3 pts**

Select all sentences that correctly define errors.

**Correct!** Errors committed on training records are training errors.**Correct!** Generalization error are expected errors of the model on previously unseen records. Low training errors guarantee low generalization errors.

Generalization errors can be reduced by increasing the model complexity.

**Question 9****3 / 3 pts**

True or False? The tree induction algorithm C4.5 is suitable for large dataset.

 True False

C4.5 is suitable for small data set.

**Correct!****Question 10****3 / 3 pts**

ID	Device OS	Carrier	Device Color	Earlier Adopter (Class)
1	Android	T-Mobile	White	Y
2	Android	AT&T	Gray	Y
3	Android	AT&T	Gray	Y
4	Android	AT&T	Black	Y
5	Android	AT&T	Pink	Y
6	Android	AT&T	Pink	Y
7	IOS	AT&T	White	Y
8	IOS	AT&T	White	Y
9	IOS	AT&T	Gray	Y
10	IOS	Verizon	Black	Y
11	Android	T-Mobile	Black	N
12	Android	T-Mobile	Pink	N
13	Android	T-Mobile	Gray	N
14	Android	Verizon	Pink	N
15	IOS	Verizon	White	N
16	IOS	Verizon	White	N
17	IOS	Verizon	Gray	N
18	IOS	Verizon	Gray	N
19	IOS	Verizon	Gray	N
20	IOS	Verizon	Black	N

Compute the Gini index for Device Color attribute using multiway split

Correct!

0.491

Incorrect Answers

Between 0.49 and 0.4915

### Question 11

3 / 3 pts

Review the table below.

ID	Device OS	Carrier	Device Color	Earlier Adopter (Class)
1	Android	T-Mobile	White	Y
2	Android	AT&T	Gray	Y
3	Android	AT&T	Gray	Y
4	Android	AT&T	Black	Y
5	Android	AT&T	Pink	Y
6	Android	AT&T	Pink	Y
7	IOS	AT&T	White	Y
8	IOS	AT&T	White	Y
9	IOS	AT&T	Gray	Y
10	IOS	Verizon	Black	Y
11	Android	T-Mobile	Black	N
12	Android	T-Mobile	Pink	N
13	Android	T-Mobile	Gray	N
14	Android	Verizon	Pink	N
15	IOS	Verizon	White	N
16	IOS	Verizon	White	N
17	IOS	Verizon	Gray	N
18	IOS	Verizon	Gray	N
19	IOS	Verizon	Gray	N
20	IOS	Verizon	Black	N

Based on the table, which of the following statements are correct about the decision tree model?

**Correct!**

- 
- The ID is the best attribute to predict Earlier Adopter
  - The Carrier is the best attribute to predict Earlier Adopter

---

  - The Device OS is the best attribute to predict Earlier Adopter

---

  - The Device Color is the best attribute to predict Earlier Adopter

**Question 12**

**3 / 3 pts**

Review the table below.

ID	Device OS	Carrier	Device Color	Earlier Adopter (Class)
1	Android	T-Mobile	White	Y
2	Android	AT&T	Gray	Y
3	Android	AT&T	Gray	Y
4	Android	AT&T	Black	Y
5	Android	AT&T	Pink	Y
6	Android	AT&T	Pink	Y
7	iOS	AT&T	White	Y
8	iOS	AT&T	White	Y
9	iOS	AT&T	Gray	Y
10	iOS	Verizon	Black	Y
11	Android	T-Mobile	Black	N
12	Android	T-Mobile	Pink	N
13	Android	T-Mobile	Gray	N
14	Android	Verizon	Pink	N
15	iOS	Verizon	White	N
16	iOS	Verizon	White	N
17	iOS	Verizon	Gray	N
18	iOS	Verizon	Gray	N
19	iOS	Verizon	Gray	N
20	iOS	Verizon	Black	N

Based on the table, which of the following statements are correct about a decision tree model?

Correct!

- Gini index for each ID value is 0.

Correct!

- Gini for Android in Device OS is the same with Gini for iOS in Device OS.

Correct!

- Gini index for the carrier is 0.1625.

Correct!

- Gini index for the overall collection of the example is 0.5.

**Question 13****3 / 3 pts**

Review the following table.

Instance	Attribute 1	Attribute 2	Attribute 3	class
1	T	T	1	Y
2	T	T	6	Y
3	T	F	5	N
4	F	F	4	Y
5	F	T	7	N
6	F	T	3	N
7	F	F	8	N
8	T	F	7	Y
9	F	T	5	N

What is the best split between Attribute 1 and Attribute 2 according to the Gini index? What is the Gini index of the split attribute?

**Correct!**

- Attribute 1 is the best split and Gini index is 0.3444
- Attribute 1 is the best split and Gini index is 0.4889
- Attribute 2 is the best split and Gini index is 0.3444
- Attribute 2 is the best split and Gini index is 0.4889

**Question 14****3 / 3 pts**

Select all classification techniques.

**Correct!**

- Decision tree

**Correct!**

- Naïive Bayes

- Clustering

**Correct!**

- Neural Networks

Decision tree, Naïve Bayes and Neural Network are commonly used classification techniques.

**Question 15****3 / 3 pts**

A good classification model has:

- Low training error
- Low estimated generalization error
- A and B are correct

We need both low training error and low generalization error.

**Correct!**

- None of the above

**Question 16****3 / 3 pts**

Review the following table.

Instance	Attribute 1	Attribute 2	Attribute 3	class
1	T	T	1	Y
2	T	T	6	Y
3	T	F	5	N
4	F	F	4	Y
5	F	T	7	N
6	F	T	3	N
7	F	F	8	N
8	T	F	7	Y
9	F	T	5	N

Select the correct information gain pair of Attribute 1 and Attribute 2.

**Correct!**

- Attribute 1- 0.229, Attribute 2 – 0.007
- Attribute 1- 0.991, Attribute 2 – 0.007
- Attribute 1- 0.007, Attribute 2 – 0.229
- Attribute 1- 0.229, Attribute 2 – 0.991

Quiz Score: **47** out of 48

# Midterm

**Due** Oct 28 at 11:59pm      **Points** 48      **Questions** 16

**Available** Oct 26 at 12am - Oct 29 at 1am 3 days

**Time Limit** 100 Minutes

## Attempt History

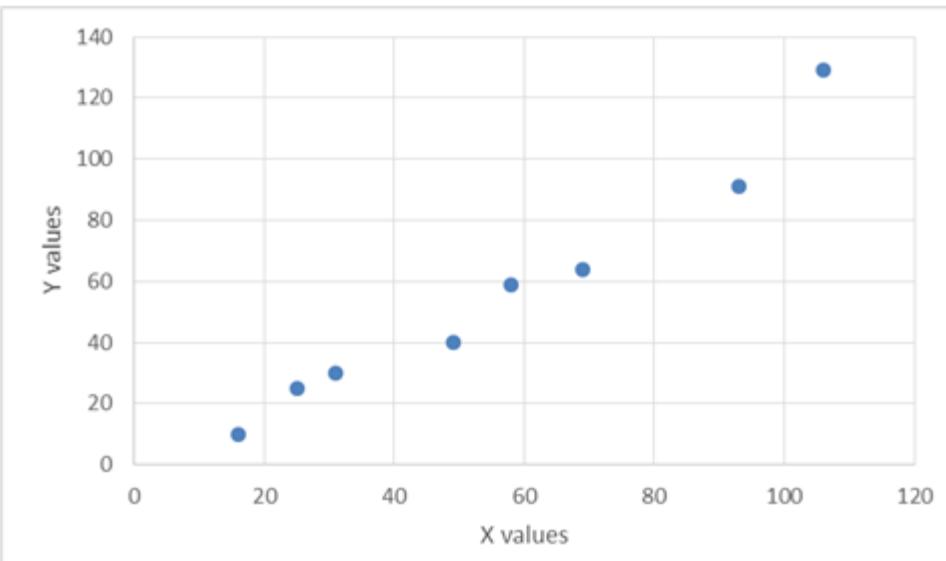
	<b>Attempt</b>	<b>Time</b>	<b>Score</b>
<b>LATEST</b>	<u><a href="#">Attempt 1</a></u>	10 minutes	45 out of 48

Score for this quiz: **45** out of 48

Submitted Oct 27 at 3:15pm

This attempt took 10 minutes.

Question 1			3 / 3 pts
	X	Y	
A	16	10	
B	25	25	
C	31	30	
D	49	40	
E	58	59	
F	69	64	
G	93	91	
H	106	129	



Compute cosine similarity between X and Y

Correct!

0.965

Incorrect Answers

Between 0.95 and 1

## Question 2

3 / 3 pts

Select all correct sentences that define correlation.

Correct!



Correlation -1 means that the vectors have a perfect negative linear relationship.



Correct. If the correlation is -1, the vectors follow the perfect negative linear correlation.



Correlation 1 means that the vectors have a perfect negative linear relationship.



Correlation 0 means that the vectors have a perfect positive linear relationship.

**Question 3****3 / 3 pts**

Select all unsupervised learning applications

**Correct!**

- Marketing and sales promotion using Association Rule

Association Rule Mining is an unsupervised learning method.

**Correct!**

- Document clustering

Clustering is an unsupervised learning method.

- Face recognition

- Fraud detection in credit card transactions

**Question 4****3 / 3 pts**

Tid	Refund	Marital Status	Taxable Income	Cheat
1	Yes	Single	125K	No
2	No	Married	100K	No
3	No	Single	70K	No
4	Yes	Married	120K	No
5	No	Divorced	95K	Yes
6	No	Married	60K	No
7	Yes	Divorced	220K	No
8	No	Single	85K	Yes
9	No	Married	75K	No
10	No	Single	90K	Yes

Select all possible columns which can be set as the class columns.

Tid

**Correct!**

Refund

**Correct!**

Marital Status

**Correct!**

Taxable Income

**Correct!**

Cheat

## Question 5

3 / 3 pts

Select all correct sentences which make up the clustering.

K-Nearest Neighbor is a typical clustering technique

**Correct!**

Association rule discovery is a typical technique of clustering

Clustering techniques do not require any parameters

A, B A, B, and C**Question 6****3 / 3 pts**

Select the correct sentence or sentences that define cosine and correlation measures.

 The range of values that are possible for the cosine measure is [0, 1] If two objects have a cosine similarity of 1, they are identical If X and Y have a mean of 0,  $\text{corr}(X,Y) = \cos(X,Y)$  A, B A, B and C**Correct!****Question 7****3 / 3 pts**

Select all sentences that correctly define errors.

**Correct!** Errors committed on training records are training errors.**Correct!**  
Generalization error are expected errors of the model on previously unseen records. Low training errors guarantee low generalization errors.

Generalization errors can be reduced by increasing the model complexity.

**Question 8****3 / 3 pts**

Review the following table.

Instance	Attribute 1	Attribute 2	Attribute 3	class
1	T	T	1	Y
2	T	T	6	Y
3	T	F	5	N
4	F	F	4	Y
5	F	T	7	N
6	F	T	3	N
7	F	F	8	N
8	T	F	7	Y
9	F	T	5	N

When is the best split for Attribute 3 according to the information gain?

**Correct!**

- The split point equal to 2.0
- The split point equal to 3.5
- The split point equal to 5.5
- The split point equal to 6.5
- The split point equal to 7.5

**Question 9****3 / 3 pts**

Review the following table.

Instance	Attribute 1	Attribute 2	Attribute 3	class
1	T	T	1	Y
2	T	T	6	Y
3	T	F	5	N
4	F	F	4	Y
5	F	T	7	N
6	F	T	3	N
7	F	F	8	N
8	T	F	7	Y
9	F	T	5	N

Select the best split among attribute 1, attribute 2, and attribute 3, according to the information gain.

**Correct!** Attribute 1 Attribute 2 Attribute 3**Question 10****3 / 3 pts**

Review the following table.

Instance	Attribute 1	Attribute 2	Attribute 3	class
1	T	T	1	Y
2	T	T	6	Y
3	T	F	5	N
4	F	F	4	Y
5	F	T	7	N
6	F	T	3	N
7	F	F	8	N
8	T	F	7	Y
9	F	T	5	N

What is the best split between Attribute 1 and Attribute 2 according to the classification rate and what is the classification error of the split attribute?

**Correct!**

- Attribute 1 is the best split, 2/9 error rate
- Attribute 1 is the best split, 4/9 error rate
- Attribute 2 is the best split, 2/9 error rate
- Attribute 2 is the best split, 4/9 error rate

### Question 11

3 / 3 pts

A good classification model has:

- Low training error
- Low estimated generalization error

**Correct!**

- A and B are correct

We need both low training error and low generalization error.

- None of the above

**Question 12****3 / 3 pts**

ID	Device OS	Carrier	Device Color	Earlier Adopter (Class)
1	Android	T-Mobile	White	Y
2	Android	AT&T	Gray	Y
3	Android	AT&T	Gray	Y
4	Android	AT&T	Black	Y
5	Android	AT&T	Pink	Y
6	Android	AT&T	Pink	Y
7	IOS	AT&T	White	Y
8	IOS	AT&T	White	Y
9	IOS	AT&T	Gray	Y
10	IOS	Verizon	Black	Y
11	Android	T-Mobile	Black	N
12	Android	T-Mobile	Pink	N
13	Android	T-Mobile	Gray	N
14	Android	Verizon	Pink	N
15	IOS	Verizon	White	N
16	IOS	Verizon	White	N
17	IOS	Verizon	Gray	N
18	IOS	Verizon	Gray	N
19	IOS	Verizon	Gray	N
20	IOS	Verizon	Black	N

Compute the Gini index for Device Color attribute using multiway split

**Correct!**

0.491

**Incorrect Answers**

Between 0.49 and 0.4915

**Question 13****3 / 3 pts**

Select all statements that correctly address underfitting and overfitting.

**Correct!** Underfitting can be reduced by reducing bias error. Overfitting can be reduced by reducing bias error. Underfitting can be reduced by reducing variance error.**Correct!** Overfitting can be reduced by reducing variance error.**Question 14****3 / 3 pts**

Review the following table.

Ground Truth	apple									
Predict	apple	pear	pear	apple	apple	pear	pear	pear	apple	apple

Compute the accuracy as a fraction

**Correct!**

0.5

**Incorrect Answers**

0.5 (with margin: 0)

**Question 15****3 / 3 pts**

Select the statements below that are computed as part of the performance metric.

**Correct!**

Precision can be computed using True Positive and False Positive.

Recall can be computed using True Positive and False Positive.

**Correct!**

F1 can be computed using True Positive and False Positive, and False Negative.

Accuracy can be computed using True Positive, False Positive, and False Negative.

**Question 16****0 / 3 pts**

Select all sentences that are correct about cross-validation.

**Correct!**

Leave-one-out is a type of cross-validation.

**You Answered**

Cross-validation guarantees reduction in test error.

K-fold cross-validation requires K repetitions of the same train set on K different test sets..

**Correct Answer**

In cross validation the same training record can be selected multiple times.

**Quiz Score: 45 out of 48**

# Midterm

**Due** Oct 28 at 11:59pm      **Points** 48      **Questions** 16

**Available** Oct 26 at 12am - Oct 29 at 1am 3 days

**Time Limit** 100 Minutes

## Attempt History

	<b>Attempt</b>	<b>Time</b>	<b>Score</b>
<b>LATEST</b>	<a href="#"><u>Attempt 1</u></a>	40 minutes	44.25 out of 48

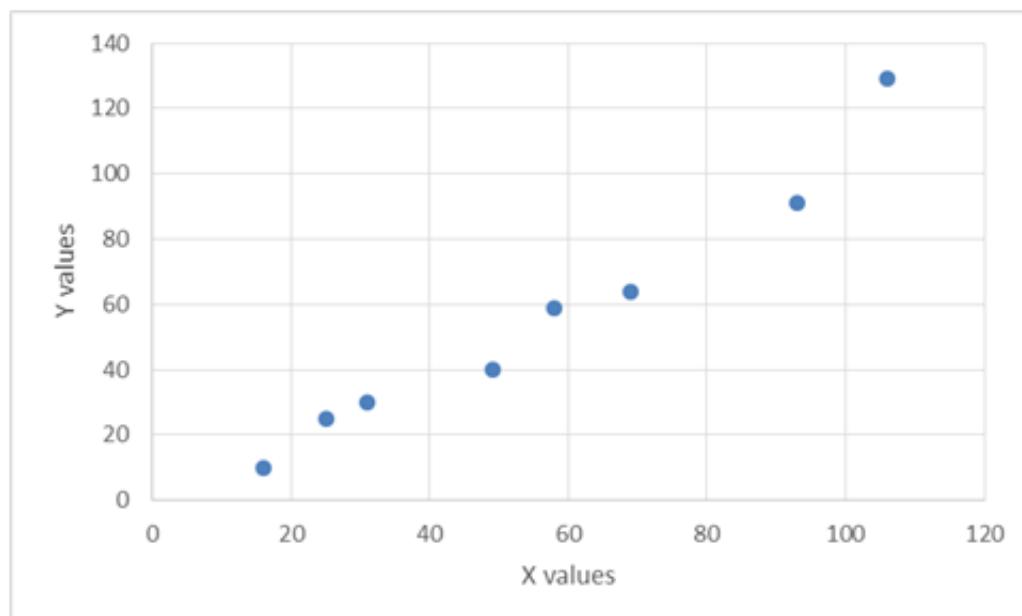
Score for this quiz: **44.25** out of 48

Submitted Oct 26 at 8:23pm

This attempt took 40 minutes.

Question 1	3 / 3 pts

	X	Y
A	16	10
B	25	25
C	31	30
D	49	40
E	58	59
F	69	64
G	93	91
H	106	129



Select the closes answer for the correlation between X and Y

0.5 -0.5 -1 1 0**Correct!****Question 2****3 / 3 pts**

Select all correct sentences that define correlation.

**Correct!**

Correlation -1 means that the vectors have a perfect negative linear relationship.

Correct. If the correlation is -1, the vectors follow the perfect negative linear correlation.



Correlation 1 means that the vectors have a perfect negative linear relationship.



Correlation 0 means that the vectors have a perfect positive linear relationship.

**Question 3****2.25 / 3 pts**

Tid	Refund	Marital Status	Taxable Income	Cheat
1	Yes	Single	125K	No
2	No	Married	100K	No
3	No	Single	70K	No
4	Yes	Married	120K	No
5	No	Divorced	95K	Yes
6	No	Married	60K	No
7	Yes	Divorced	220K	No
8	No	Single	85K	Yes
9	No	Married	75K	No
10	No	Single	90K	Yes

Select all possible columns which can be set as the class columns.

 Tid**Correct!** Refund**Correct!** Marital Status**Correct Answer** Taxable Income**Correct!** Cheat**Question 4****3 / 3 pts**

Select all unsupervised learning applications

Correct!

- Marketing and sales promotion using Association Rule

Association Rule Mining is an unsupervised learning method.

Correct!

- Document clustering

Clustering is an unsupervised learning method.

- Face recognition

- Fraud detection in credit card transactions

## Question 5

3 / 3 pts

Select all correct sentences which make up the clustering.

- K-Nearest Neighbor is a typical clustering technique

- Association rule discovery is a typical technique of clustering

- Clustering techniques do not require any parameters

- A, B

- A, B, and C

**Question 6**

0 / 3 pts

Select the correct sentence or sentences that define cosine and correlation measures.



The range of values that are possible for the cosine measure is [0, 1]



If two objects have a cosine similarity of 1, they are identical



If  $X$  and  $Y$  have a mean of 0,  $\text{corr}(X, Y) = \cos(X, Y)$



A, B



A, B and C

**Correct Answer****You Answered****Question 7**

3 / 3 pts

Select all classification techniques.

**Correct!** Decision tree**Correct!** Naïive Bayes**Correct!** Clustering Neural Networks

Decision tree, Naïve Bayes and Neural Network are commonly used classification techniques.

**Question 8****3 / 3 pts**

Review the table below.

ID	Device OS	Carrier	Device Color	Earlier Adopter (Class)
1	Android	T-Mobile	White	Y
2	Android	AT&T	Gray	Y
3	Android	AT&T	Gray	Y
4	Android	AT&T	Black	Y
5	Android	AT&T	Pink	Y
6	Android	AT&T	Pink	Y
7	IOS	AT&T	White	Y
8	IOS	AT&T	White	Y
9	IOS	AT&T	Gray	Y
10	IOS	Verizon	Black	Y
11	Android	T-Mobile	Black	N
12	Android	T-Mobile	Pink	N
13	Android	T-Mobile	Gray	N
14	Android	Verizon	Pink	N
15	IOS	Verizon	White	N
16	IOS	Verizon	White	N
17	IOS	Verizon	Gray	N
18	IOS	Verizon	Gray	N
19	IOS	Verizon	Gray	N
20	IOS	Verizon	Black	N

Based on the table, which of the following statements are correct about a decision tree model?

Correct!

- Gini index for each ID value is 0.

Correct!



Gini for Android in Device OS is the same with Gini for IOS in Device OS.

Correct!

- Gini index for the carrier is 0.1625.

Correct!

- Gini index for the overall collection of the example is 0.5.

### Question 9

3 / 3 pts

Review the table below.

ID	Device OS	Carrier	Device Color	Earlier Adopter (Class)
1	Android	T-Mobile	White	Y
2	Android	AT&T	Gray	Y
3	Android	AT&T	Gray	Y
4	Android	AT&T	Black	Y
5	Android	AT&T	Pink	Y
6	Android	AT&T	Pink	Y
7	IOS	AT&T	White	Y
8	IOS	AT&T	White	Y
9	IOS	AT&T	Gray	Y
10	IOS	Verizon	Black	Y
11	Android	T-Mobile	Black	N
12	Android	T-Mobile	Pink	N
13	Android	T-Mobile	Gray	N
14	Android	Verizon	Pink	N
15	IOS	Verizon	White	N
16	IOS	Verizon	White	N
17	IOS	Verizon	Gray	N
18	IOS	Verizon	Gray	N
19	IOS	Verizon	Gray	N
20	IOS	Verizon	Black	N

Based on the table, which of the following statements are correct about the decision tree model?

- 
- The ID is the best attribute to predict Earlier Adopter
  - The Carrier is the best attribute to predict Earlier Adopter
  - The Device OS is the best attribute to predict Earlier Adopter
  - The Device Color is the best attribute to predict Earlier Adopter

Correct!

**Question 10****3 / 3 pts**

Review the following table.

Instance	Attribute 1	Attribute 2	Attribute 3	class
1	T	T	1	Y
2	T	T	6	Y
3	T	F	5	N
4	F	F	4	Y
5	F	T	7	N
6	F	T	3	N
7	F	F	8	N
8	T	F	7	Y
9	F	T	5	N

What is the best split between Attribute 1 and Attribute 2 according to the classification rate and what is the classification error of the split attribute?

**Correct!**

- Attribute 1 is the best split, 2/9 error rate
- Attribute 1 is the best split, 4/9 error rate
- Attribute 2 is the best split, 2/9 error rate
- Attribute 2 is the best split, 4/9 error rate

**Question 11****3 / 3 pts**

ID	Device OS	Carrier	Device Color	Earlier Adopter (Class)
1	Android	T-Mobile	White	Y
2	Android	AT&T	Gray	Y
3	Android	AT&T	Gray	Y
4	Android	AT&T	Black	Y
5	Android	AT&T	Pink	Y
6	Android	AT&T	Pink	Y
7	IOS	AT&T	White	Y
8	IOS	AT&T	White	Y
9	IOS	AT&T	Gray	Y
10	IOS	Verizon	Black	Y
11	Android	T-Mobile	Black	N
12	Android	T-Mobile	Pink	N
13	Android	T-Mobile	Gray	N
14	Android	Verizon	Pink	N
15	IOS	Verizon	White	N
16	IOS	Verizon	White	N
17	IOS	Verizon	Gray	N
18	IOS	Verizon	Gray	N
19	IOS	Verizon	Gray	N
20	IOS	Verizon	Black	N

Compute the Gini index for Device Color attribute using multiway split

**Correct!**

0.491

**Correct Answers**

Between 0.49 and 0.4915

**Question 12**

3 / 3 pts

Review the following table.

Ground Truth	apple									
Predict	apple	pear	pear	apple	apple	pear	pear	pear	apple	apple

Compute F1 for 'pear' class

**Correct!** 0**Correct Answers**

0 (with margin: 0)

**Question 13**

3 / 3 pts

Review the table below.

Ground Truth	apple									
Predict	apple	pear	pear	apple	apple	pear	pear	pear	apple	apple

Select the correct precision and recall for 'apple' class.

- 
- Precision – 0.5 and recall – 0.5

**Correct!**

- Precision – 1 and recall – 0.5
- Precision – 0.5 and recall – 1
- Precision – 1 and recall – 1

**Question 14****3 / 3 pts**

Review the following table.

Instance	Attribute 1	Attribute 2	Attribute 3	class
1	T	T	1	Y
2	T	T	6	Y
3	T	F	5	N
4	F	F	4	Y
5	F	T	7	N
6	F	T	3	N
7	F	F	8	N
8	T	F	7	Y
9	F	T	5	N

Select the correct information gain pair of Attribute 1 and Attribute 2.

**Correct!**

- Attribute 1- 0.229, Attribute 2 – 0.007

- Attribute 1- 0.991, Attribute 2 – 0.007
- Attribute 1- 0.007, Attribute 2 – 0.229
- Attribute 1- 0.229, Attribute 2 – 0.991

**Question 15****3 / 3 pts**

Select all sentences that correctly define errors.

**Correct!**

- Prediction errors consist of bias, variance, and irreducible error.

**Correct!**

- Bias error results in high training error.

- Variance error results in high training error.

**Correct!**

- Irreducible error cannot be optimized.

**Question 16****3 / 3 pts**

Select Methods for good model evaluation.

**Correct!**

- Maximize values of both precision and recall

we need to maximize both precision and recall

- Maximize accuracy
- Maximize precision
- Minimize recall

Quiz Score: **44.25** out of 48

# Module 2 Practice Quiz

**Due** No due date    **Points** 10    **Questions** 10  
**Time Limit** None    **Allowed Attempts** Unlimited

## Instructions

10 questions!

[Take the Quiz Again](#)

## Attempt History

	<b>Attempt</b>	<b>Time</b>	<b>Score</b>
LATEST	<a href="#"><u>Attempt 1</u></a>	less than 1 minute	0.58 out of 10

Submitted Oct 28 at 12:52pm

### Question 1

0.25 / 1 pts

Review the table below.

ID	Device OS	Carrier	Device Color	Earlier Adopter (Class)
1	Android	T-Mobile	White	Y
2	Android	AT&T	Gray	Y
3	Android	AT&T	Gray	Y
4	Android	AT&T	Black	Y
5	Android	AT&T	Pink	Y
6	Android	AT&T	Pink	Y
7	IOS	AT&T	White	Y
8	IOS	AT&T	White	Y
9	IOS	AT&T	Gray	Y
10	IOS	Verizon	Black	Y
11	Android	T-Mobile	Black	N
12	Android	T-Mobile	Pink	N
13	Android	T-Mobile	Gray	N
14	Android	Verizon	Pink	N
15	IOS	Verizon	White	N
16	IOS	Verizon	White	N
17	IOS	Verizon	Gray	N
18	IOS	Verizon	Gray	N
19	IOS	Verizon	Gray	N
20	IOS	Verizon	Black	N

Based on the table, which of the following builds the decision tree model?

Correct Answer

- Gini index for each ID value is 0.

Correct Answer

Gini for Android in Device OS is the same with Gini for IOS in Device OS.

Correct!

- Gini index for the carrier is 0.1625.

Correct Answer

- Gini index for the overall collection of the example is 0.5.

Question 2

0 / 1 pts

Review the table below.

ID	Device OS	Carrier	Device Color	Earlier Adopter (Class)
1	Android	T-Mobile	White	Y
2	Android	AT&T	Gray	Y
3	Android	AT&T	Gray	Y
4	Android	AT&T	Black	Y
5	Android	AT&T	Pink	Y
6	Android	AT&T	Pink	Y
7	IOS	AT&T	White	Y
8	IOS	AT&T	White	Y
9	IOS	AT&T	Gray	Y
10	IOS	Verizon	Black	Y
11	Android	T-Mobile	Black	N
12	Android	T-Mobile	Pink	N
13	Android	T-Mobile	Gray	N
14	Android	Verizon	Pink	N
15	IOS	Verizon	White	N
16	IOS	Verizon	White	N
17	IOS	Verizon	Gray	N
18	IOS	Verizon	Gray	N
19	IOS	Verizon	Gray	N
20	IOS	Verizon	Black	N

Based on the table, which of the following builds the decision tree model?

Correct Answer

- The ID is the best attribute to predict Earlier Adopter

DU Answered

- The Carrier is the best attribute to predict Earlier Adopter

- The Device OS is the best attribute to predict Earlier Adopter

- The Device Color is the best attribute to predict Earlier Adopter

Inanswered

Question 3

0 / 1 pts

ID	Device OS	Carrier	Device Color	Earlier Adopter (Class)
1	Android	T-Mobile	White	Y
2	Android	AT&T	Gray	Y
3	Android	AT&T	Gray	Y
4	Android	AT&T	Black	Y
5	Android	AT&T	Pink	Y
6	Android	AT&T	Pink	Y
7	IOS	AT&T	White	Y
8	IOS	AT&T	White	Y
9	IOS	AT&T	Gray	Y
10	IOS	Verizon	Black	Y
11	Android	T-Mobile	Black	N
12	Android	T-Mobile	Pink	N
13	Android	T-Mobile	Gray	N
14	Android	Verizon	Pink	N
15	IOS	Verizon	White	N
16	IOS	Verizon	White	N
17	IOS	Verizon	Gray	N
18	IOS	Verizon	Gray	N
19	IOS	Verizon	Gray	N
20	IOS	Verizon	Black	N

Compute the Gini index for Device Color attribute using multiway split

DU Answered

orrect Answers

Between 0.49 and 0.4915

Inanswered

#### Question 4

0 / 1 pts

Select all sentences that correctly define errors.

orrect Answer

- Errors committed on training records are training errors.

**Correct Answer**

Generalization error are expected errors of the model on previously unseen records.

Low training errors guarantee low generalization errors.

Generalization errors can be reduced by increasing the model complexity.

**Question 5****0.33 / 1 pts**

Select all sentences that correctly define errors.

**Correct!**

Prediction errors consist of bias, variance, and irreducible error.

**Correct Answer**

Errors due to bias are primarily caused by models that underfit.

**Correct Answer**

Errors due to variance are caused by models that overfit.

Irreducible errors are caused by complex models.

**Inanswered****Question 6****0 / 1 pts**

Review the following table.

Ground Truth	apple									
Predict	apple	pear	pear	apple	apple	pear	pear	pear	apple	apple

Compute F1 for 'pear' class

**You Answered****Correct Answers**

0 (with margin: 0)

**Inanswered****Question 7**

0 / 1 pts

Review the following table.

Ground Truth	apple									
Predict	apple	pear	pear	apple	apple	pear	pear	pear	apple	apple

Compute the accuracy

**You Answered****Correct Answers**

0.5 (with margin: 0)

**Inanswered****Question 8**

0 / 1 pts

Review the following table.

i	ii	iii	# of instances	
			+	-
T	T	T	5	0
F	T	T	0	20
T	F	T	20	0
F	F	T	0	5
T	T	F	0	0
F	T	F	25	0
T	F	F	0	0
F	F	F	0	25

According to the classification error rate, which attribute would be chosen as the first splitting attribute?

Correct Answer

'i'

'ii'

'iii'

anyone

Inanswered

### Question 9

0 / 1 pts

Review the following table.

i	ii	iii	# of instances	
			+	-
T	T	T	5	0
F	T	T	0	20
T	F	T	20	0
F	F	T	0	5
T	T	F	0	0
F	T	F	25	0
T	F	F	0	0
F	F	F	0	25

After the first splitting and 'i' = F child node, which attribute would be chosen as the second splitting attribute?

'i'

'ii'

'iii'

anyone

Correct Answer

In answered

### Question 10

0 / 1 pts

Review the following table.

I	II	III	# of instances	
			+	-
T	T	T	5	0
F	T	T	0	20
T	F	T	20	0
F	F	T	0	5
T	T	F	0	0
F	T	F	25	0
T	F	F	0	0
F	F	F	0	25

Compute the accuracy by the resulting decision tree.

You Answered

Correct Answers

0.2 (with margin: 0)

20 (with margin: 0)

# Knowledge Check: Classification Issues

---

<b>Due</b> No due date	<b>Points</b> 4	<b>Questions</b> 4
<b>Time Limit</b> None	<b>Allowed Attempts</b> Unlimited	

---

## Instructions

4 questions!

[Take the Quiz Again](#)

## Attempt History

	<b>Attempt</b>	<b>Time</b>	<b>Score</b>
<b>LATEST</b>	<a href="#"><u>Attempt 1</u></a>	1 minute	1.5 out of 4

---

Submitted Oct 24 at 2:15am

**Question 1** 0 / 1 pts

A good classification model has:

A) Low training error

B) Low generalization error

B) Low generalization error alone is not sufficient for a good classification model.

**Answered**

**Correct Answer**

C) A and B are correct

D) None of the above

---

**Question 2****1 / 1 pts**

Choose the best method for good model evaluation.

**Correct!**

- Maximize values of both precision and recall

we need to maximize both precision and recall

- Maximize accuracy
- Maximize precision
- Minimize recall

**Question 3****0.5 / 1 pts**

Select all that are true about estimating generalization errors.

**Correct!**

- Generalization error is always minimized by minimizing training error



Given two models of similar generalization errors, one should prefer the simpler model over the more complex model

Complex model may not perform well on unseen data.

**Correct Answer**

- Complex models can accidentally fit errors in data
- All of the above

**Question 4****0 / 1 pts**

Which of the following are true regarding bias and variance tradeoff?

**You Answered**

- Underfitting is caused by low variance alone.

We need to consider also bias when evaluating the model.

**Correct Answer**

- Simple model tend to have High variance and low bias.
- Complex models overfit because of low variance
- None of the above.

# Knowledge Check: Introduction to Classification Tasks

<b>Due</b> No due date	<b>Points</b> 4	<b>Questions</b> 4
<b>Time Limit</b> None	<b>Allowed Attempts</b> Unlimited	

## Instructions

4 questions!

[Take the Quiz Again](#)

## Attempt History

	<b>Attempt</b>	<b>Time</b>	<b>Score</b>
<b>LATEST</b>	<a href="#">Attempt 1</a>	less than 1 minute	0 out of 4

Submitted Oct 28 at 12:54pm

In answered

**Question 1** 0 / 1 pts

Select all techniques used to measure node impurity.

Correct Answer

Gini Index

average

Correct Answer

Entropy

split

average and split do not give you information about node's impurity

Inanswered

**Question 2**

0 / 1 pts

True or False? The tree induction algorithm C4.5 is suitable for large dataset.

 True False

Correct Answer

Inanswered

**Question 3**

0 / 1 pts

Which are not issues of decision trees?

 Specifying the attribute test condition. Determining the best split. Determine when to stop splitting. Splitting on discrete attributes. All of the above

Correct Answer

Inanswered

**Question 4**

0 / 1 pts

Select all that is true about best split of node of a decision tree.

Correct Answer

Non-homogeneous nodes have high degree of impurity

Correct Answer

Nodes with homogeneous class distribution are preferred

Correct Answer

Homogeneous nodes have low degree of impurity.

None of the above

# Knowledge Check: Introduction to Classification

**Due** No due date    **Points** 2    **Questions** 2  
**Time Limit** None    **Allowed Attempts** Unlimited

## Instructions

2 questions!

[Take the Quiz Again](#)

## Attempt History

	<b>Attempt</b>	<b>Time</b>	<b>Score</b>
LATEST	<a href="#">Attempt 1</a>	less than 1 minute	0 out of 2

Submitted Oct 28 at 12:54pm

In answered	<b>Question 1</b>	0 / 1 pts
<p>True or False? A test set is used to find a model for class attribute as a function of the values of other attributes</p> <p><input type="radio"/> True</p> <p><input checked="" type="radio"/> False</p>		

In answered	<b>Question 2</b>	0 / 1 pts
<p>Select all classification techniques.</p>		

**orrect Answer** Decision tree**orrect Answer** Rule-based systems Clustering**orrect Answer** Neural Networks

Clustering is an unsupervised learning method.

# Knowledge Check: Review of Initial Data Exploration Techniques

<b>Due</b> No due date	<b>Points</b> 2	<b>Questions</b> 2
<b>Time Limit</b> None	<b>Allowed Attempts</b> Unlimited	

## Instructions

2 questions.

LO: Review and summarize data exploration techniques for use in initial data analysis

[Take the Quiz Again](#)

## Attempt History

	<b>Attempt</b>	<b>Time</b>	<b>Score</b>
LATEST	<a href="#">Attempt 1</a>	less than 1 minute	0 out of 2

Submitted Oct 28 at 1:02pm

In answered	<b>Question 1</b>	0 / 1 pts
	Select all correct sentences that make up the similarity property.	
Correct Answer	<input type="checkbox"/> If $p = q$ , similarity, $s(p, q)$ is one (or maximum similarity)	
	<input type="checkbox"/> Similarity is always in the range 0 to 1	
Correct Answer	<input type="checkbox"/> Similarity is usually in the range -1 to 1	

**Correct Answer**

Cosine and Jaccard-Coefficient are typical similarity measurement methods

In case of non-negative inputs, similarity is in the range 0 to 1. So it is not ALWAYS in range -1 to 1. However, correlation is in range -1 to 1

**Inanswered****Question 2**

0 / 1 pts

Select all correct sentences that define correlation.



Correlation 0 means that the vectors have a perfect positive linear relationship.



Correlation -1 means that the vectors have a perfect negative linear relationship.



Correlation 1 means that the vectors have a perfect negative linear relationship.

**Correct Answer**

# Module 1 Practice Quiz

**Due** No due date    **Points** 10    **Questions** 10  
**Time Limit** None    **Allowed Attempts** Unlimited

## Instructions

10 questions!

LO:

- 1.1 Explain the history and purpose of data mining across multiple disciplines
- 1.2 Differentiate what is and what is not data mining
- 1.3 Describe different data mining tasks
- 1.4 Recognize attributes of data needed for data mining
- 1.5 Review and summarize data exploration techniques for use in initial data analysis

[Take the Quiz Again](#)

## Attempt History

	Attempt	Time	Score
LATEST	<a href="#">Attempt 1</a>	less than 1 minute	0 out of 10

Submitted Oct 28 at 12:54pm

In answered	<b>Question 1</b>	0 / 1 pts
	Select all directly relevant fields to data mining that have influenced the breakthroughs in the field of data mining.	
Correct Answer	<input type="checkbox"/> Machine Learning / Pattern Recognition	

**orrect Answer**

- 
- Statistics / Artificial Intelligence

**orrect Answer**

- 
- Database System

- 
- Computer Network

**Inanswered****Question 2**

0 / 1 pts

What factor or factors have influenced the rise of the use of data mining breakthroughs?



Services like Youtube become popular and allowed for the collection of large amounts of data.



Sensor data collection with pervasive manner became more common



Hardware became cheaper and more powerful



A, B



A, B and C

**orrect Answer****Inanswered****Question 3**

0 / 1 pts

Select all correct pairs of data attributes type and the appropriate example.

**orrect Answer**

- 
- Nominal – hair colors

**orrect Answer**

- 
- Interval – centigrade scale

**orrect Answer**

- Ratio – Degree Kelvin or absolute

**orrect Answer**

- Ordinal – order of finishing a race

**Inanswered****Question 4****0 / 1 pts**

Select all correct sentences that make up the classification

- Classification is an unsupervised learning method

- Feature extraction is a mandatory preprocessing

**orrect Answer**

- Classification is a supervised learning method

**orrect Answer**

- Spam filtering is a typical classification application

- Dimensionality reduction is a typical classification technique

**Inanswered****Question 5****0 / 1 pts**

Select all possible use-cases of clustering

- Credit card fraud detection

- Market segmentation

- Summarize news

**orrect Answer**

- A, B, and C

- A, B

**Inanswered****Question 6****0 / 1 pts**

Select all sentences that are true.

K-Nearest Neighbor is a typical clustering technique

Association rule discovery is a typical technique for supervised learning

Clustering techniques do not require any parameters

A, B

None of the above

**Correct Answer****Inanswered****Question 7****0 / 1 pts**

Select all predictive techniques.

Regression

**Correct Answer**

Classification

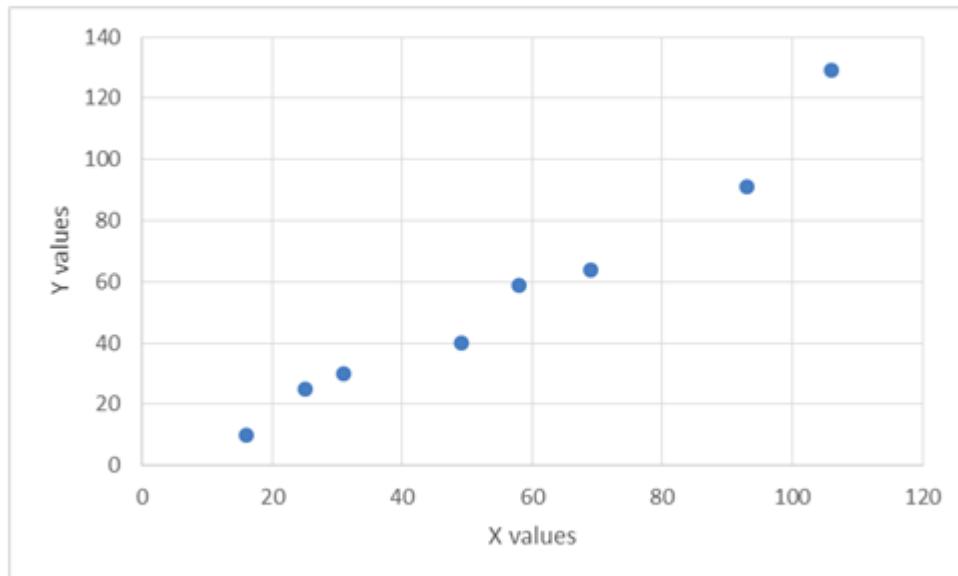
**Correct Answer**

Clustering

Dimensionality Reduction

**Inanswered****Question 8****0 / 1 pts**

	X	Y
A	16	10
B	25	25
C	31	30
D	49	40
E	58	59
F	69	64
G	93	91
H	106	129



Select the correct answer of Euclidean distance between (A and B), (D and E), and (G and H)

7.81(A and B), 39.62(D and E), 50.99(G and H)

17.49(A and B), 21.02(D and E), 50.99(G and H)

- 24(A and B), 28(D and E), 51(G and H)

- None

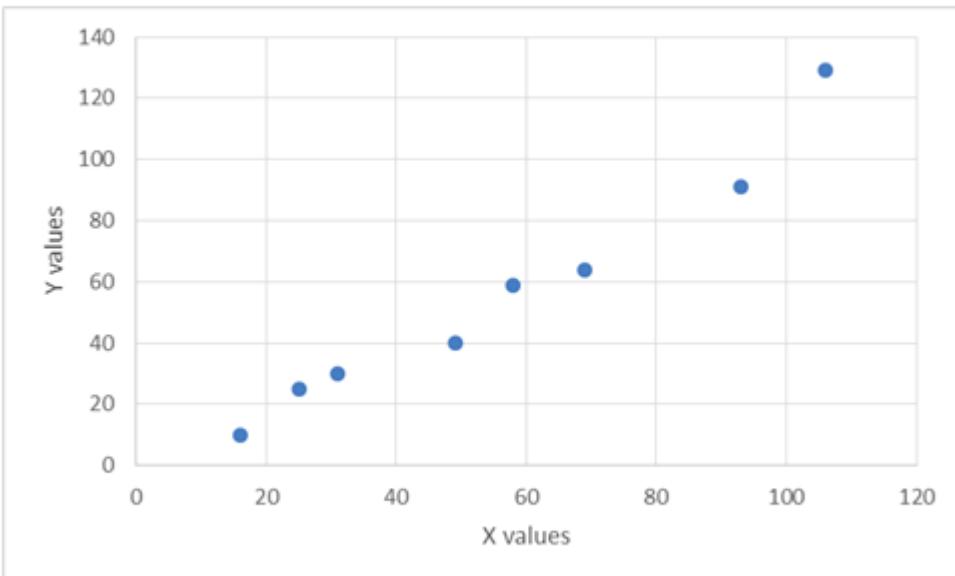
Correct Answer

Inanswered

### Question 9

0 / 1 pts

	X	Y
A	16	10
B	25	25
C	31	30
D	49	40
E	58	59
F	69	64
G	93	91
H	106	129



Compute cosine similarity between X and Y

DU Answered

Correct Answers

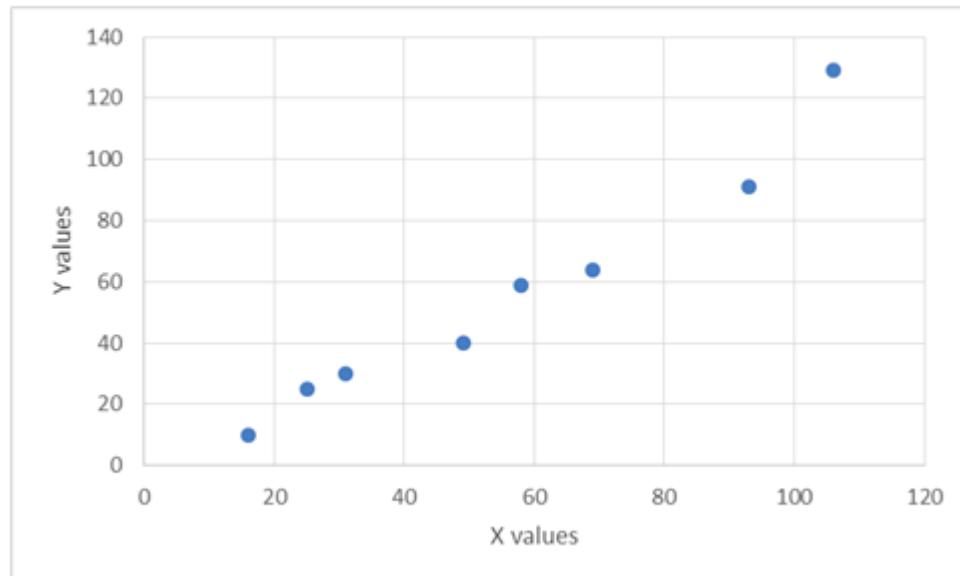
Between 0.99 and 0.992

Inanswered

### Question 10

0 / 1 pts

	X	Y
A	16	10
B	25	25
C	31	30
D	49	40
E	58	59
F	69	64
G	93	91
H	106	129



Select the most similar correlation between X and Y

0.5

-0.5

-1

**Correct Answer** 1 0

# Knowledge Check: Data Attributes Needed for Data Mining

**Due** No due date    **Points** 2    **Questions** 2  
**Time Limit** None    **Allowed Attempts** Unlimited

## Instructions

2 questions!

LO: Recognize attributes of data needed for data mining

[Take the Quiz Again](#)

## Attempt History

	Attempt	Time	Score
LATEST	<a href="#">Attempt 1</a>	1 minute	0 out of 2

Submitted Oct 28 at 1:01pm

In answered	Question 1	0 / 1 pts
	Select all correct sentences that define attributes.	
	<input type="checkbox"/> The continuous attributes are often represented as integer variables.	
	<input type="checkbox"/> The binary attributes are discrete attributes.	
Correct Answer	<input checked="" type="checkbox"/>	The discrete attribute has only a finite or countably infinite set of values.

Unanswered

**Question 2**

0 / 1 pts

Select from the list below the attributes that could contribute to data quality problems.

**Correct Answer** Outliers**Correct Answer** Missing values**Correct Answer** Noise**Correct Answer** Duplicate Data

Noise refers to the modification of original values. Outliers are data objects with characteristics that are considerably different than most of the other data objects in the data set. Missing values and duplicate data are intuitively critical to data quality.

# Knowledge Check: Introduction to Data Mining Tasks

<b>Due</b> No due date	<b>Points</b> 2	<b>Questions</b> 2
<b>Time Limit</b> None	<b>Allowed Attempts</b> Unlimited	

## Instructions

2 questions!

LO: Describe different data mining tasks

[Take the Quiz Again](#)

## Attempt History

	<b>Attempt</b>	<b>Time</b>	<b>Score</b>
LATEST	<a href="#">Attempt 1</a>	less than 1 minute	0 out of 2

Submitted Oct 28 at 1pm

In answered	<b>Question 1</b>	0 / 1 pts
Select all supervised learning methodologies		
<input type="checkbox"/> Clustering		
<input type="checkbox"/> Dimensionality Reduction		
Correct Answer	<input type="checkbox"/> Classification or categorization	
Correct Answer	<input type="checkbox"/> Regression	

Unanswered

**Question 2**

0 / 1 pts

Select all unsupervised learning applications

Correct Answer

- Marketing and sales promotion using Association Rule

Correct Answer

- Document clustering

- Face recognition

- Fraud detection in credit card transactions

# Knowledge Check: History and Purpose of Data Mining

**Due** No due date    **Points** 2    **Questions** 2  
**Time Limit** None    **Allowed Attempts** Unlimited

## Instructions

2 questions!

LO: Explain the history and purpose of data mining across multiple disciplines

[Take the Quiz Again](#)

## Attempt History

	Attempt	Time	Score
LATEST	<a href="#">Attempt 1</a>	less than 1 minute	0 out of 2

Submitted Oct 28 at 12:59pm

In answered	<b>Question 1</b>	0 / 1 pts
	Discovering a student name in the class roster is an example of data mining function.	
	<input type="radio"/> True	
Correct Answer	<input checked="" type="radio"/> False	

Since the student name in the class roster can be obtained by looking up, the student name is already known information (or knowledge).

Unanswered

**Question 2**

0 / 1 pts

Big data is required for data mining.

 True False

Correct Answer

Although the increase in collected data size necessitates data mining, there is no specific data size needed to perform data mining.

# Midterm

**Due** Oct 28 at 11:59pm      **Points** 48      **Questions** 16

**Available** Oct 26 at 12am - Oct 29 at 1am 3 days

**Time Limit** 100 Minutes

## Attempt History

	<b>Attempt</b>	<b>Time</b>	<b>Score</b>
<b>LATEST</b>	<a href="#"><u>Attempt 1</u></a>	87 minutes	44 out of 48

Score for this quiz: **44** out of 48

Submitted Oct 27 at 7:37pm

This attempt took 87 minutes.

<b>Question 1</b>	<b>3 / 3 pts</b>
Calculate Jaccard similar between X and Y  X = 0101010001  Y = 0100011000	
<b>Correct!</b>  0.4	
<b>Correct Answers</b> 0.4 (with margin: 0)	

<b>Question 2</b>	<b>3 / 3 pts</b>

Tid	Refund	Marital Status	Taxable Income	Cheat
1	Yes	Single	125K	No
2	No	Married	100K	No
3	No	Single	70K	No
4	Yes	Married	120K	No
5	No	Divorced	95K	Yes
6	No	Married	60K	No
7	Yes	Divorced	220K	No
8	No	Single	85K	Yes
9	No	Married	75K	No
10	No	Single	90K	Yes

Select all possible columns which can be set as the class columns.

Tid

Correct!

Refund

Correct!

Marital Status

Correct!

Taxable Income

Correct!

Cheat

### Question 3

2 / 3 pts

Select all predictive techniques.

Anomaly Detection

Correct Answer

**Correct!** Regression**Correct!** Classification Clustering Dimensionality Reduction**Question 4****3 / 3 pts**

Select all unsupervised learning applications

**Correct!** Marketing and sales promotion using Association Rule

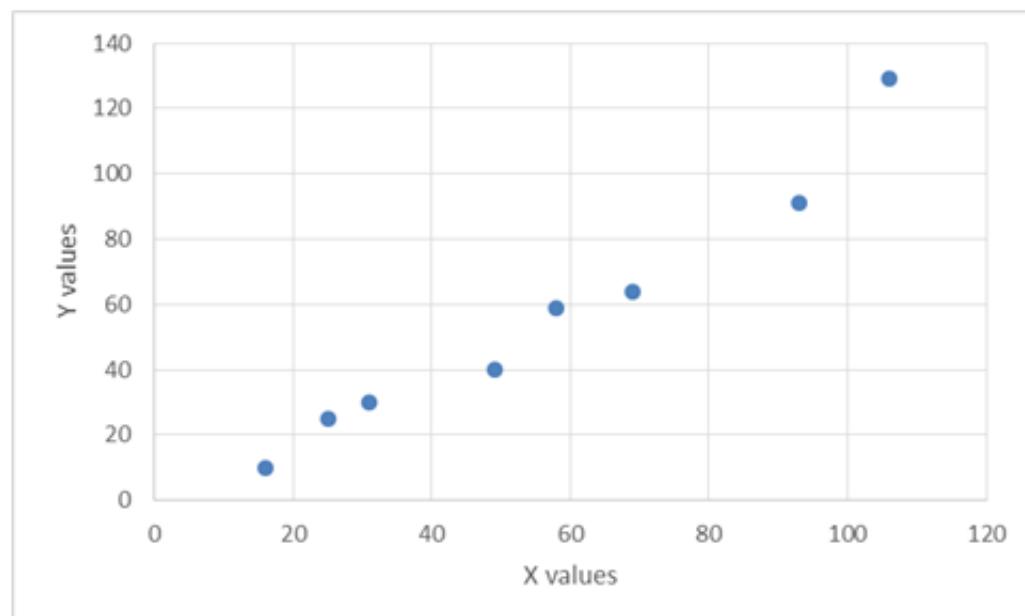
Association Rule Mining is an unsupervised learning method.

**Correct!** Document clustering

Clustering is an unsupervised learning method.

 Face recognition Fraud detection in credit card transactions**Question 5****3 / 3 pts**

	X	Y
A	16	10
B	25	25
C	31	30
D	49	40
E	58	59
F	69	64
G	93	91
H	106	129



Select the closes answer for the correlation between X and Y

**Correct!** 0.5 -0.5 -1 1 0**Question 6****3 / 3 pts**

Select all correct sentences that define correlation.

**Correct!**

Correlation -1 means that the vectors have a perfect negative linear relationship.

Correct. If the correlation is -1, the vectors follow the perfect negative linear correlation.



Correlation 1 means that the vectors have a perfect negative linear relationship.



Correlation 0 means that the vectors have a perfect positive linear relationship.

**Question 7****3 / 3 pts**

Select all sentences that correctly define errors.

**Correct!**

- Prediction errors consist of bias, variance, and irreducible error.

**Correct!**

- Bias error results in high training error.

- Variance error results in high training error.

**Correct!**

- Irreducible error cannot be optimized.

**Question 8****0 / 3 pts**

Select all sentences that are correct about cross-validation.

**Correct!**

- Leave-one-out is a type of cross-validation.

**You Answered**

- Cross-validation guarantees reduction in test error.

- 

K-fold cross-validation requires K repetitions of the same train set on K different test sets..

**Correct Answer**

- 

In cross validation the same training record can be selected multiple times.

**Question 9****3 / 3 pts**

Select all statements that correctly address underfitting and overfitting.

**Correct!**

- Underfitting can be reduced by reducing bias error.
- Overfitting can be reduced by reducing bias error.
- Underfitting can be reduced by reducing variance error.

**Correct!**

- Overfitting can be reduced by reducing variance error.

**Question 10****3 / 3 pts**

Review the table below.

Ground Truth	apple									
Predict	apple	pear	pear	apple	apple	pear	pear	pear	apple	apple

Select the correct precision and recall for 'apple' class.

**Correct!**

- Precision – 0.5 and recall – 0.5
- Precision – 1 and recall – 0.5
- Precision – 0.5 and recall – 1

- Precision – 1 and recall – 1

**Question 11**

3 / 3 pts

Review the following table.

Ground Truth	apple									
Predict	apple	pear	pear	apple	apple	pear	pear	pear	apple	apple

Compute the accuracy as a fraction

Correct!

0.5

Correct Answers

0.5 (with margin: 0)

**Question 12**

3 / 3 pts

Review the table below.

ID	Device OS	Carrier	Device Color	Earlier Adopter (Class)
1	Android	T-Mobile	White	Y
2	Android	AT&T	Gray	Y
3	Android	AT&T	Gray	Y
4	Android	AT&T	Black	Y
5	Android	AT&T	Pink	Y
6	Android	AT&T	Pink	Y
7	IOS	AT&T	White	Y
8	IOS	AT&T	White	Y
9	IOS	AT&T	Gray	Y
10	IOS	Verizon	Black	Y
11	Android	T-Mobile	Black	N
12	Android	T-Mobile	Pink	N
13	Android	T-Mobile	Gray	N
14	Android	Verizon	Pink	N
15	IOS	Verizon	White	N
16	IOS	Verizon	White	N
17	IOS	Verizon	Gray	N
18	IOS	Verizon	Gray	N
19	IOS	Verizon	Gray	N
20	IOS	Verizon	Black	N

Based on the table, which of the following statements are correct about the decision tree model?

Correct!

- 
- The ID is the best attribute to predict Earlier Adopter
  - The Carrier is the best attribute to predict Earlier Adopter
  - The Device OS is the best attribute to predict Earlier Adopter
  - The Device Color is the best attribute to predict Earlier Adopter
-

**Question 13**

3 / 3 pts

Review the following table.

Ground Truth	apple									
Predict	apple	pear	pear	apple	apple	pear	pear	pear	apple	apple

Compute F1 for 'pear' class

Correct!

Correct Answers

0 (with margin: 0)

**Question 14**

3 / 3 pts

Select the statements below that are computed as part of the performance metric.

Correct!

 Precision can be computed using True Positive and False Positive. Recall can be computed using True Positive and False Positive.

Correct!

F1 can be computed using True Positive and False Positive, and False Negative.



Accuracy can be computed using True Positive, False Positive, and False Negative.

### Question 15

3 / 3 pts

Review the following table.

Instance	Attribute 1	Attribute 2	Attribute 3	class
1	T	T	1	Y
2	T	T	6	Y
3	T	F	5	N
4	F	F	4	Y
5	F	T	7	N
6	F	T	3	N
7	F	F	8	N
8	T	F	7	Y
9	F	T	5	N

What is the best split between Attribute 1 and Attribute 2 according to the Gini index? What is the Gini index of the split attribute?

Correct!

- Attribute 1 is the best split and Gini index is 0.3444

- Attribute 1 is the best split and Gini index is 0.4889
- Attribute 2 is the best split and Gini index is 0.3444
- Attribute 2 is the best split and Gini index is 0.4889

**Question 16****3 / 3 pts**

Select all classification techniques.

**Correct!**

- Decision tree

**Correct!**

- Naïive Bayes

- Clustering

**Correct!**

- Neural Networks

Decision tree, Naïive Bayes and Neural Network are commonly used classification techniques.

**Quiz Score: 44 out of 48**

# Midterm

Due Oct 28 at 11:59pm

Points 48

Questions 16

Available Oct 26 at 12am - Oct 29 at 1am 3 days

Time Limit 100 Minutes

## Attempt History

	Attempt	Time	Score
LATEST	<a href="#">Attempt 1</a>	71 minutes	47 out of 48

Score for this quiz: **47** out of 48

Submitted Oct 27 at 7:47pm

This attempt took 71 minutes.

Question 1					3 / 3 pts
<hr/>					
Tid	Refund	Marital Status	Taxable Income	Cheat	
1	Yes	Single	125K	No	
2	No	Married	100K	No	
3	No	Single	70K	No	
4	Yes	Married	120K	No	
5	No	Divorced	95K	Yes	
6	No	Married	60K	No	
7	Yes	Divorced	220K	No	
8	No	Single	85K	Yes	
9	No	Married	75K	No	
10	No	Single	90K	Yes	

Select all possible columns which can be set as the class columns.

Tid

**Correct!** Refund**Correct!** Marital Status**Correct!** Taxable Income**Correct!** Cheat**Question 2****3 / 3 pts**

Select the correct sentence or sentences that define cosine and correlation measures.



The range of values that are possible for the cosine measure is [0, 1]



If two objects have a cosine similarity of 1, they are identical



If X and Y have a mean of 0,  $\text{corr}(X,Y) = \cos(X,Y)$



A, B



A, B and C

**Question 3****3 / 3 pts**

Calculate Jaccard similar between X and Y

X = 0101010001

Y = 0100011000

**Correct!**

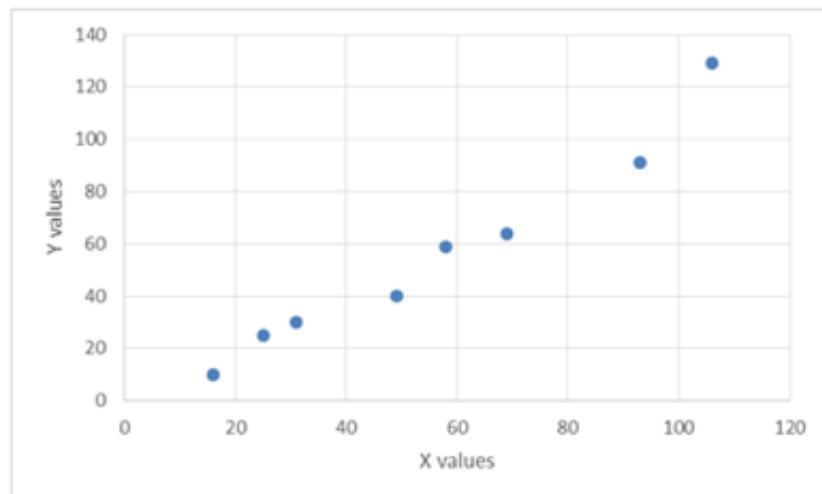
0.4

**Correct Answers**

0.4 (with margin: 0)

**Question 4****3 / 3 pts**

	X	Y
A	16	10
B	25	25
C	31	30
D	49	40
E	58	59
F	69	64
G	93	91
H	106	129



Compute cosine similarity between X and Y

Correct!

0.991

Correct Answers

Between 0.95 and 1

## Question 5

3 / 3 pts

Select all unsupervised learning applications

Correct!

Marketing and sales promotion using Association Rule

Association Rule Mining is an unsupervised learning method.

Correct!

Document clustering

Clustering is an unsupervised learning method.

Face recognition

- Fraud detection in credit card transactions

**Question 6****2 / 3 pts**

Select all predictive techniques.

**Incorrect Answer**

- Anomaly Detection

**Correct!**

- Regression

**Correct!**

- Classification

- Clustering

- Dimensionality Reduction

**Question 7****3 / 3 pts**

Review the following table.

Instance	Attribute 1	Attribute 2	Attribute 3	class
1	T	T	1	Y
2	T	T	6	Y
3	T	F	5	N
4	F	F	4	Y
5	F	T	7	N
6	F	T	3	N
7	F	F	8	N
8	T	F	7	Y
9	F	T	5	N

What is the best split between Attribute 1 and Attribute 2 according to the classification rate and what is the classification error of the split attribute?

Correct!

- Attribute 1 is the best split, 2/9 error rate
- Attribute 1 is the best split, 4/9 error rate
- Attribute 2 is the best split, 2/9 error rate
- Attribute 2 is the best split, 4/9 error rate

## Question 8

3 / 3 pts

Select all sentences that are correct about cross-validation.

Correct!

- Leave-one-out is a type of cross-validation.

- Cross-validation guarantees reduction in test error.
- K-fold cross-validation requires K repetitions of the same train set on K different test sets..

Correct!

- In cross validation the same training record can be selected multiple times.

### Question 9

3 / 3 pts

Select Methods for good model evaluation.

Correct!

- Maximize values of both precision and recall
  - we need to maximize both precision and recall
- Maximize accuracy
- Maximize precision
- Minimize recall

### Question 10

3 / 3 pts

True or False? A test set is used to find a model for class attribute as a function of the values of other attributes

True False

Correct. A test set is used to determine the accuracy of the model.

### Question 11

3 / 3 pts

Review the table below.

ID	Device OS	Carrier	Device Color	Earlier Adopter (Class)
1	Android	T-Mobile	White	Y
2	Android	AT&T	Gray	Y
3	Android	AT&T	Gray	Y
4	Android	AT&T	Black	Y
5	Android	AT&T	Pink	Y
6	Android	AT&T	Pink	Y
7	IOS	AT&T	White	Y
8	IOS	AT&T	White	Y
9	IOS	AT&T	Gray	Y
10	IOS	Verizon	Black	Y
11	Android	T-Mobile	Black	N
12	Android	T-Mobile	Pink	N
13	Android	T-Mobile	Gray	N
14	Android	Verizon	Pink	N
15	IOS	Verizon	White	N
16	IOS	Verizon	White	N
17	IOS	Verizon	Gray	N
18	IOS	Verizon	Gray	N
19	IOS	Verizon	Gray	N
20	IOS	Verizon	Black	N

Based on the table, which of the following statements are correct about a decision tree model?

Correct!

- Gini index for each ID value is 0.

Correct!

- Gini for Android in Device OS is the same with Gini for IOS in Device OS.

Correct!

- Gini index for the carrier is 0.1625.

Correct!

- Gini index for the overall collection of the example is 0.5.

### Question 12

3 / 3 pts

Review the table below.

ID	Device OS	Carrier	Device Color	Earlier Adopter (Class)
1	Android	T-Mobile	White	Y
2	Android	AT&T	Gray	Y
3	Android	AT&T	Gray	Y
4	Android	AT&T	Black	Y
5	Android	AT&T	Pink	Y
6	Android	AT&T	Pink	Y
7	IOS	AT&T	White	Y
8	IOS	AT&T	White	Y
9	IOS	AT&T	Gray	Y
10	IOS	Verizon	Black	Y
11	Android	T-Mobile	Black	N
12	Android	T-Mobile	Pink	N
13	Android	T-Mobile	Gray	N
14	Android	Verizon	Pink	N
15	IOS	Verizon	White	N
16	IOS	Verizon	White	N
17	IOS	Verizon	Gray	N
18	IOS	Verizon	Gray	N
19	IOS	Verizon	Gray	N
20	IOS	Verizon	Black	N

Based on the table, which of the following statements are correct about the decision tree model?

- 
- The ID is the best attribute to predict Earlier Adopter
  - The Carrier is the best attribute to predict Earlier Adopter
  - The Device OS is the best attribute to predict Earlier Adopter
  - The Device Color is the best attribute to predict Earlier Adopter

Correct!

**Question 13**

3 / 3 pts

Review the following table.

Instance	Attribute 1	Attribute 2	Attribute 3	class
1	T	T	1	Y
2	T	T	6	Y
3	T	F	5	N
4	F	F	4	Y
5	F	T	7	N
6	F	T	3	N
7	F	F	8	N
8	T	F	7	Y
9	F	T	5	N

Select the best split among attribute 1, attribute 2, and attribute 3, according to the information gain.

---

**Correct!**

---

 Attribute 1

---

 Attribute 2

---

 Attribute 3

---

**Question 14**

3 / 3 pts

Review the following table.

Instance	Attribute 1	Attribute 2	Attribute 3	class
1	T	T	1	Y
2	T	T	6	Y
3	T	F	5	N
4	F	F	4	Y
5	F	T	7	N
6	F	T	3	N
7	F	F	8	N
8	T	F	7	Y
9	F	T	5	N

Select the correct information gain pair of Attribute 1 and Attribute 2.

Correct!

- Attribute 1- 0.229, Attribute 2 – 0.007
- Attribute 1- 0.991, Attribute 2 – 0.007
- Attribute 1- 0.007, Attribute 2 – 0.229
- Attribute 1- 0.229, Attribute 2 – 0.991

### Question 15

3 / 3 pts

Review the following table.

Ground Truth	apple									
Predict	apple	pear	pear	apple	apple	pear	pear	pear	apple	apple

Compute the accuracy as a fraction

Correct!

0.5

Correct Answers

0.5 (with margin: 0)

### Question 16

3 / 3 pts

Select the statements below that are computed as part of the performance metric.

Correct!



Precision can be computed using True Positive and False Positive.



Recall can be computed using True Positive and False Positive.

Correct!



F1 can be computed using True Positive and False Positive, and False Negative.



Accuracy can be computed using True Positive, False Positive, and False Negative.

Quiz Score: 47 out of 48

# Midterm

**Due** Oct 28 at 11:59pm      **Points** 48      **Questions** 16

**Available** Oct 26 at 12am - Oct 29 at 1am 3 days

**Time Limit** 100 Minutes

## Attempt History

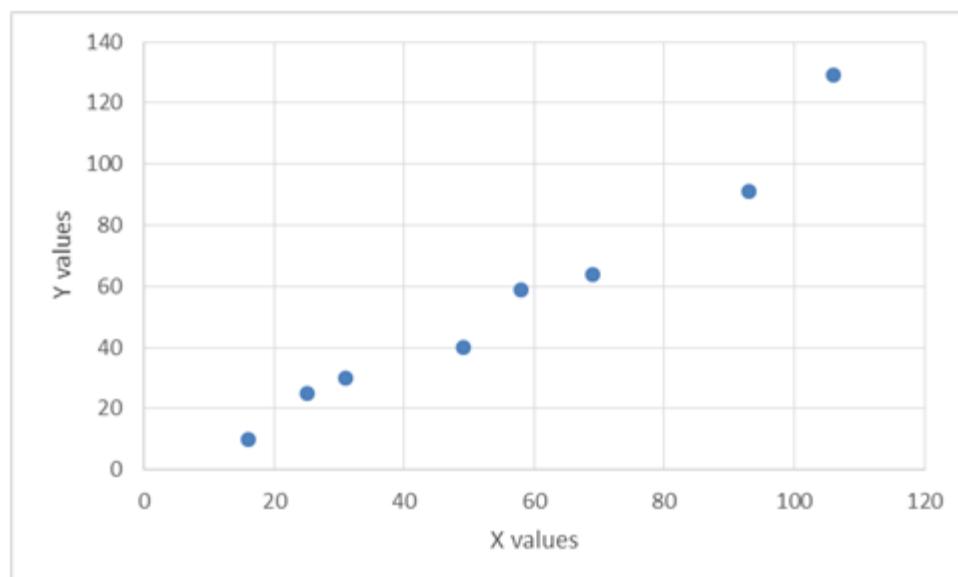
	Attempt	Time	Score
<b>LATEST</b>	<a href="#"><u>Attempt 1</u></a>	13 minutes	47 out of 48

Score for this quiz: **47** out of 48

Submitted Oct 27 at 3:23pm

This attempt took 13 minutes.

Question 1			3 / 3 pts																																				
<table border="1"><thead><tr><th></th><th>X</th><th>Y</th><th></th></tr></thead><tbody><tr><td>A</td><td>16</td><td>10</td><td></td></tr><tr><td>B</td><td>25</td><td>25</td><td></td></tr><tr><td>C</td><td>31</td><td>30</td><td></td></tr><tr><td>D</td><td>49</td><td>40</td><td></td></tr><tr><td>E</td><td>58</td><td>59</td><td></td></tr><tr><td>F</td><td>69</td><td>64</td><td></td></tr><tr><td>G</td><td>93</td><td>91</td><td></td></tr><tr><td>H</td><td>106</td><td>129</td><td></td></tr></tbody></table>					X	Y		A	16	10		B	25	25		C	31	30		D	49	40		E	58	59		F	69	64		G	93	91		H	106	129	
	X	Y																																					
A	16	10																																					
B	25	25																																					
C	31	30																																					
D	49	40																																					
E	58	59																																					
F	69	64																																					
G	93	91																																					
H	106	129																																					



Compute cosine similarity between X and Y

Correct!

0.973

Incorrect Answers

Between 0.95 and 1

## Question 2

3 / 3 pts

Select all correct sentences which make up the clustering.

- K-Nearest Neighbor is a typical clustering technique
- Association rule discovery is a typical technique of clustering
- Clustering techniques do not require any parameters
- A, B
- A, B, and C

Correct!

**Question 3****3 / 3 pts**

Select all possible use-cases of clustering

- Credit card fraud detection
- Market segmentation
- Summarize news

**Correct!**

- A, B, and C
- A, B

**Question 4****2 / 3 pts**

Select all predictive techniques.

**Incorrect Answer**

- Anomaly Detection

**Correct!**

- Regression

**Correct!**

- Classification

- Clustering

- Dimensionality Reduction

**Question 5****3 / 3 pts**

Select the correct sentence or sentences that define cosine and correlation

measures.

The range of values that are possible for the cosine measure is [0, 1]

If two objects have a cosine similarity of 1, they are identical

**Correct!**

If  $X$  and  $Y$  have a mean of 0,  $\text{corr}(X,Y) = \cos(X,Y)$

A, B

A, B and C

### Question 6

3 / 3 pts

Tid	Refund	Marital Status	Taxable Income	Cheat
1	Yes	Single	125K	No
2	No	Married	100K	No
3	No	Single	70K	No
4	Yes	Married	120K	No
5	No	Divorced	95K	Yes
6	No	Married	60K	No
7	Yes	Divorced	220K	No
8	No	Single	85K	Yes
9	No	Married	75K	No
10	No	Single	90K	Yes

Select all possible columns which can be set as the class columns.

Tid

**Correct!**

Refund

**Correct!**

Marital Status

**Correct!** Taxable Income**Correct!** Cheat**Question 7****3 / 3 pts**

Select the statements below that are computed as part of the performance metric.

**Correct!** Precision can be computed using True Positive and False Positive. Recall can be computed using True Positive and False Positive.**Correct!** F1 can be computed using True Positive and False Positive, and False Negative. Accuracy can be computed using True Positive, False Positive, and False Negative.**Question 8****3 / 3 pts**

Select all sentences that correctly define errors.

**Correct!** Errors committed on training records are training errors.**Correct!** Generalization error are expected errors of the model on previously unseen records. Low training errors guarantee low generalization errors.

Generalization errors can be reduced by increasing the model complexity.

**Question 9****3 / 3 pts**

True or False? The tree induction algorithm C4.5 is suitable for large dataset.

 True False

C4.5 is suitable for small data set.

**Correct!****Question 10****3 / 3 pts**

ID	Device OS	Carrier	Device Color	Earlier Adopter (Class)
1	Android	T-Mobile	White	Y
2	Android	AT&T	Gray	Y
3	Android	AT&T	Gray	Y
4	Android	AT&T	Black	Y
5	Android	AT&T	Pink	Y
6	Android	AT&T	Pink	Y
7	IOS	AT&T	White	Y
8	IOS	AT&T	White	Y
9	IOS	AT&T	Gray	Y
10	IOS	Verizon	Black	Y
11	Android	T-Mobile	Black	N
12	Android	T-Mobile	Pink	N
13	Android	T-Mobile	Gray	N
14	Android	Verizon	Pink	N
15	IOS	Verizon	White	N
16	IOS	Verizon	White	N
17	IOS	Verizon	Gray	N
18	IOS	Verizon	Gray	N
19	IOS	Verizon	Gray	N
20	IOS	Verizon	Black	N

Compute the Gini index for Device Color attribute using multiway split

Correct!

0.491

Incorrect Answers

Between 0.49 and 0.4915

### Question 11

3 / 3 pts

Review the table below.

ID	Device OS	Carrier	Device Color	Earlier Adopter (Class)
1	Android	T-Mobile	White	Y
2	Android	AT&T	Gray	Y
3	Android	AT&T	Gray	Y
4	Android	AT&T	Black	Y
5	Android	AT&T	Pink	Y
6	Android	AT&T	Pink	Y
7	IOS	AT&T	White	Y
8	IOS	AT&T	White	Y
9	IOS	AT&T	Gray	Y
10	IOS	Verizon	Black	Y
11	Android	T-Mobile	Black	N
12	Android	T-Mobile	Pink	N
13	Android	T-Mobile	Gray	N
14	Android	Verizon	Pink	N
15	IOS	Verizon	White	N
16	IOS	Verizon	White	N
17	IOS	Verizon	Gray	N
18	IOS	Verizon	Gray	N
19	IOS	Verizon	Gray	N
20	IOS	Verizon	Black	N

Based on the table, which of the following statements are correct about the decision tree model?

**Correct!**

- 
- The ID is the best attribute to predict Earlier Adopter
  - The Carrier is the best attribute to predict Earlier Adopter

---

  - The Device OS is the best attribute to predict Earlier Adopter

---

  - The Device Color is the best attribute to predict Earlier Adopter

**Question 12**

**3 / 3 pts**

Review the table below.

ID	Device OS	Carrier	Device Color	Earlier Adopter (Class)
1	Android	T-Mobile	White	Y
2	Android	AT&T	Gray	Y
3	Android	AT&T	Gray	Y
4	Android	AT&T	Black	Y
5	Android	AT&T	Pink	Y
6	Android	AT&T	Pink	Y
7	IOS	AT&T	White	Y
8	IOS	AT&T	White	Y
9	IOS	AT&T	Gray	Y
10	IOS	Verizon	Black	Y
11	Android	T-Mobile	Black	N
12	Android	T-Mobile	Pink	N
13	Android	T-Mobile	Gray	N
14	Android	Verizon	Pink	N
15	IOS	Verizon	White	N
16	IOS	Verizon	White	N
17	IOS	Verizon	Gray	N
18	IOS	Verizon	Gray	N
19	IOS	Verizon	Gray	N
20	IOS	Verizon	Black	N

Based on the table, which of the following statements are correct about a decision tree model?

Correct!

- Gini index for each ID value is 0.

Correct!

- Gini for Android in Device OS is the same with Gini for IOS in Device OS.

Correct!

- Gini index for the carrier is 0.1625.

Correct!

- Gini index for the overall collection of the example is 0.5.

**Question 13****3 / 3 pts**

Review the following table.

Instance	Attribute 1	Attribute 2	Attribute 3	class
1	T	T	1	Y
2	T	T	6	Y
3	T	F	5	N
4	F	F	4	Y
5	F	T	7	N
6	F	T	3	N
7	F	F	8	N
8	T	F	7	Y
9	F	T	5	N

What is the best split between Attribute 1 and Attribute 2 according to the Gini index? What is the Gini index of the split attribute?

**Correct!**

- Attribute 1 is the best split and Gini index is 0.3444
- Attribute 1 is the best split and Gini index is 0.4889
- Attribute 2 is the best split and Gini index is 0.3444
- Attribute 2 is the best split and Gini index is 0.4889

**Question 14****3 / 3 pts**

Select all classification techniques.

**Correct!**

- Decision tree

**Correct!**

- Naïive Bayes

- Clustering

**Correct!**

- Neural Networks

Decision tree, Naïive Bayes and Neural Network are commonly used classification techniques.

**Question 15****3 / 3 pts**

A good classification model has:

- Low training error
- Low estimated generalization error
- A and B are correct

We need both low training error and low generalization error.

**Correct!**

- None of the above

**Question 16****3 / 3 pts**

Review the following table.

Instance	Attribute 1	Attribute 2	Attribute 3	class
1	T	T	1	Y
2	T	T	6	Y
3	T	F	5	N
4	F	F	4	Y
5	F	T	7	N
6	F	T	3	N
7	F	F	8	N
8	T	F	7	Y
9	F	T	5	N

Select the correct information gain pair of Attribute 1 and Attribute 2.

**Correct!**

- Attribute 1- 0.229, Attribute 2 – 0.007
- 
- Attribute 1- 0.991, Attribute 2 – 0.007
- 
- Attribute 1- 0.007, Attribute 2 – 0.229
- 
- Attribute 1- 0.229, Attribute 2 – 0.991

Quiz Score: **47** out of 48

# Midterm

**Due** Oct 28 at 11:59pm    **Points** 48    **Questions** 16

**Available** Oct 26 at 12am - Oct 29 at 1am 3 days

**Time Limit** 100 Minutes

## Attempt History

	<b>Attempt</b>	<b>Time</b>	<b>Score</b>
<b>LATEST</b>	<a href="#"><u>Attempt 1</u></a>	10 minutes	45 out of 48

Score for this quiz: **45** out of 48

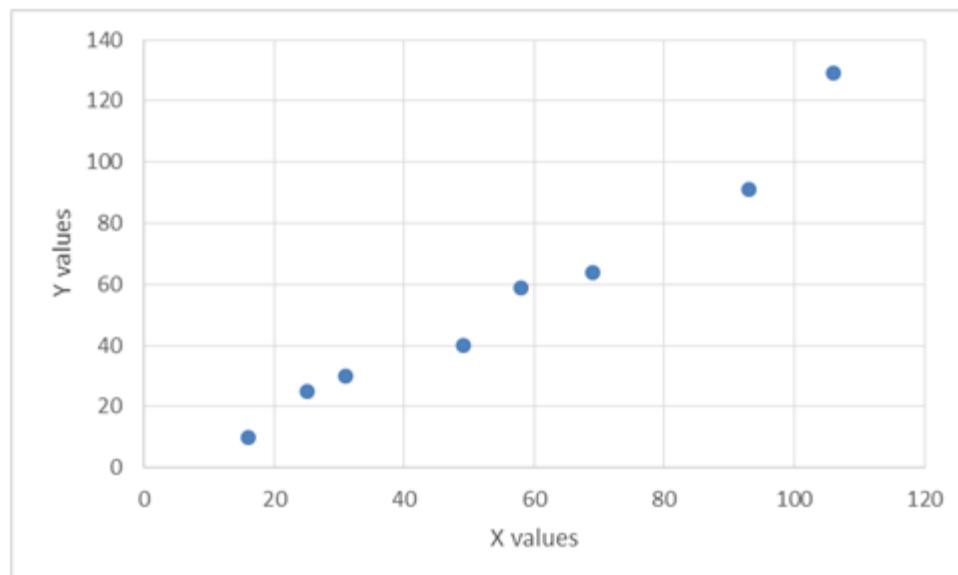
Submitted Oct 27 at 3:15pm

This attempt took 10 minutes.

### Question 1

**3 / 3 pts**

	X	Y
A	16	10
B	25	25
C	31	30
D	49	40
E	58	59
F	69	64
G	93	91
H	106	129



Compute cosine similarity between X and Y

Correct!

0.965

Incorrect Answers

Between 0.95 and 1

## Question 2

3 / 3 pts

Select all correct sentences that define correlation.

Correct!



Correlation -1 means that the vectors have a perfect negative linear relationship.



Correct. If the correlation is -1, the vectors follow the perfect negative linear correlation.



Correlation 1 means that the vectors have a perfect negative linear relationship.



Correlation 0 means that the vectors have a perfect positive linear relationship.

**Question 3****3 / 3 pts**

Select all unsupervised learning applications

**Correct!**

- Marketing and sales promotion using Association Rule

Association Rule Mining is an unsupervised learning method.

**Correct!**

- Document clustering

Clustering is an unsupervised learning method.

- Face recognition

- Fraud detection in credit card transactions

**Question 4****3 / 3 pts**

Tid	Refund	Marital Status	Taxable Income	Cheat
1	Yes	Single	125K	No
2	No	Married	100K	No
3	No	Single	70K	No
4	Yes	Married	120K	No
5	No	Divorced	95K	Yes
6	No	Married	60K	No
7	Yes	Divorced	220K	No
8	No	Single	85K	Yes
9	No	Married	75K	No
10	No	Single	90K	Yes

Select all possible columns which can be set as the class columns.

Tid

**Correct!**

Refund

**Correct!**

Marital Status

**Correct!**

Taxable Income

**Correct!**

Cheat

## Question 5

3 / 3 pts

Select all correct sentences which make up the clustering.

K-Nearest Neighbor is a typical clustering technique

**Correct!**

Association rule discovery is a typical technique of clustering

Clustering techniques do not require any parameters

A, B A, B, and C**Question 6****3 / 3 pts**

Select the correct sentence or sentences that define cosine and correlation measures.

 The range of values that are possible for the cosine measure is [0, 1] If two objects have a cosine similarity of 1, they are identical If X and Y have a mean of 0,  $\text{corr}(X,Y) = \cos(X,Y)$  A, B A, B and C**Correct!****Question 7****3 / 3 pts**

Select all sentences that correctly define errors.

**Correct!** Errors committed on training records are training errors.**Correct!**  
Generalization error are expected errors of the model on previously unseen records. Low training errors guarantee low generalization errors.

Generalization errors can be reduced by increasing the model complexity.

**Question 8****3 / 3 pts**

Review the following table.

Instance	Attribute 1	Attribute 2	Attribute 3	class
1	T	T	1	Y
2	T	T	6	Y
3	T	F	5	N
4	F	F	4	Y
5	F	T	7	N
6	F	T	3	N
7	F	F	8	N
8	T	F	7	Y
9	F	T	5	N

When is the best split for Attribute 3 according to the information gain?

**Correct!**

- The split point equal to 2.0
- The split point equal to 3.5
- The split point equal to 5.5
- The split point equal to 6.5
- The split point equal to 7.5

**Question 9****3 / 3 pts**

Review the following table.

Instance	Attribute 1	Attribute 2	Attribute 3	class
1	T	T	1	Y
2	T	T	6	Y
3	T	F	5	N
4	F	F	4	Y
5	F	T	7	N
6	F	T	3	N
7	F	F	8	N
8	T	F	7	Y
9	F	T	5	N

Select the best split among attribute 1, attribute 2, and attribute 3, according to the information gain.

**Correct!** Attribute 1 Attribute 2 Attribute 3**Question 10****3 / 3 pts**

Review the following table.

Instance	Attribute 1	Attribute 2	Attribute 3	class
1	T	T	1	Y
2	T	T	6	Y
3	T	F	5	N
4	F	F	4	Y
5	F	T	7	N
6	F	T	3	N
7	F	F	8	N
8	T	F	7	Y
9	F	T	5	N

What is the best split between Attribute 1 and Attribute 2 according to the classification rate and what is the classification error of the split attribute?

**Correct!**

- Attribute 1 is the best split, 2/9 error rate
- Attribute 1 is the best split, 4/9 error rate
- Attribute 2 is the best split, 2/9 error rate
- Attribute 2 is the best split, 4/9 error rate

### Question 11

3 / 3 pts

A good classification model has:

- Low training error
- Low estimated generalization error

**Correct!**

- A and B are correct

We need both low training error and low generalization error.

- None of the above

**Question 12****3 / 3 pts**

ID	Device OS	Carrier	Device Color	Earlier Adopter (Class)
1	Android	T-Mobile	White	Y
2	Android	AT&T	Gray	Y
3	Android	AT&T	Gray	Y
4	Android	AT&T	Black	Y
5	Android	AT&T	Pink	Y
6	Android	AT&T	Pink	Y
7	IOS	AT&T	White	Y
8	IOS	AT&T	White	Y
9	IOS	AT&T	Gray	Y
10	IOS	Verizon	Black	Y
11	Android	T-Mobile	Black	N
12	Android	T-Mobile	Pink	N
13	Android	T-Mobile	Gray	N
14	Android	Verizon	Pink	N
15	IOS	Verizon	White	N
16	IOS	Verizon	White	N
17	IOS	Verizon	Gray	N
18	IOS	Verizon	Gray	N
19	IOS	Verizon	Gray	N
20	IOS	Verizon	Black	N

Compute the Gini index for Device Color attribute using multiway split

**Correct!**

0.491

**Incorrect Answers**

Between 0.49 and 0.4915

**Question 13****3 / 3 pts**

Select all statements that correctly address underfitting and overfitting.

**Correct!** Underfitting can be reduced by reducing bias error. Overfitting can be reduced by reducing bias error. Underfitting can be reduced by reducing variance error.**Correct!** Overfitting can be reduced by reducing variance error.**Question 14****3 / 3 pts**

Review the following table.

Ground Truth	apple									
Predict	apple	pear	pear	apple	apple	pear	pear	pear	apple	apple

Compute the accuracy as a fraction

**Correct!**

0.5

**Incorrect Answers**

0.5 (with margin: 0)

**Question 15****3 / 3 pts**

Select the statements below that are computed as part of the performance metric.

**Correct!**

Precision can be computed using True Positive and False Positive.

Recall can be computed using True Positive and False Positive.

**Correct!**

F1 can be computed using True Positive and False Positive, and False Negative.

Accuracy can be computed using True Positive, False Positive, and False Negative.

**Question 16****0 / 3 pts**

Select all sentences that are correct about cross-validation.

**Correct!**

Leave-one-out is a type of cross-validation.

**You Answered**

Cross-validation guarantees reduction in test error.

K-fold cross-validation requires K repetitions of the same train set on K different test sets..

**Correct Answer**

In cross validation the same training record can be selected multiple times.

**Quiz Score: 45 out of 48**

# Midterm

**Due** Oct 28 at 11:59pm      **Points** 48      **Questions** 16

**Available** Oct 26 at 12am - Oct 29 at 1am 3 days

**Time Limit** 100 Minutes

## Attempt History

	<b>Attempt</b>	<b>Time</b>	<b>Score</b>
<b>LATEST</b>	<a href="#"><u>Attempt 1</u></a>	40 minutes	44.25 out of 48

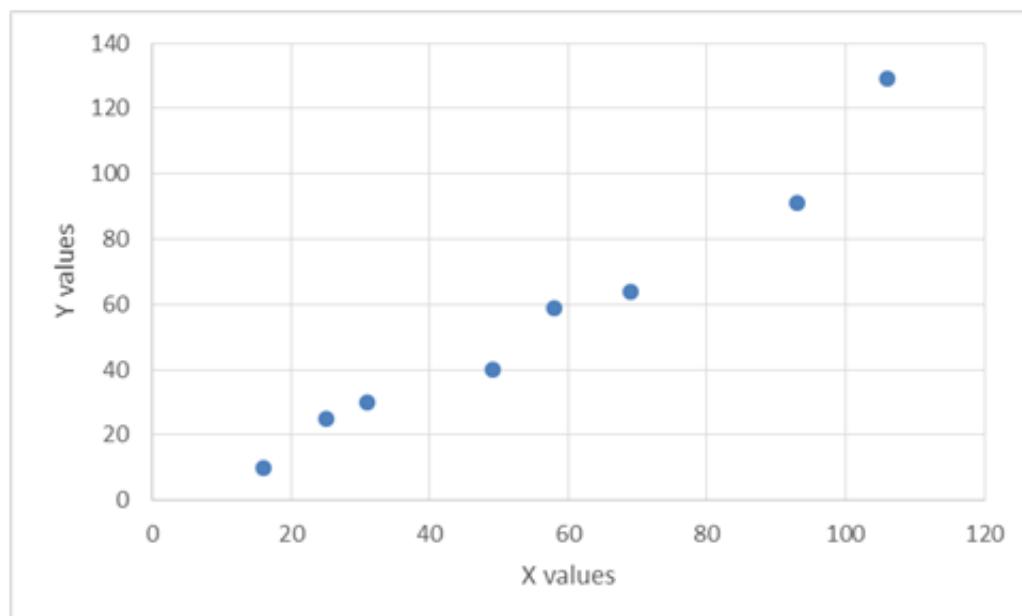
Score for this quiz: **44.25** out of 48

Submitted Oct 26 at 8:23pm

This attempt took 40 minutes.

Question 1	3 / 3 pts

	X	Y
A	16	10
B	25	25
C	31	30
D	49	40
E	58	59
F	69	64
G	93	91
H	106	129



Select the closes answer for the correlation between X and Y

0.5 -0.5 -1 1 0**Correct!****Question 2****3 / 3 pts**

Select all correct sentences that define correlation.

**Correct!**

Correlation -1 means that the vectors have a perfect negative linear relationship.

Correct. If the correlation is -1, the vectors follow the perfect negative linear correlation.



Correlation 1 means that the vectors have a perfect negative linear relationship.



Correlation 0 means that the vectors have a perfect positive linear relationship.

**Question 3****2.25 / 3 pts**

Tid	Refund	Marital Status	Taxable Income	Cheat
1	Yes	Single	125K	No
2	No	Married	100K	No
3	No	Single	70K	No
4	Yes	Married	120K	No
5	No	Divorced	95K	Yes
6	No	Married	60K	No
7	Yes	Divorced	220K	No
8	No	Single	85K	Yes
9	No	Married	75K	No
10	No	Single	90K	Yes

Select all possible columns which can be set as the class columns.

 Tid**Correct!** Refund**Correct!** Marital Status**Correct Answer** Taxable Income**Correct!** Cheat**Question 4****3 / 3 pts**

Select all unsupervised learning applications

Correct!

- Marketing and sales promotion using Association Rule

Association Rule Mining is an unsupervised learning method.

Correct!

- Document clustering

Clustering is an unsupervised learning method.

- Face recognition

- Fraud detection in credit card transactions

## Question 5

3 / 3 pts

Select all correct sentences which make up the clustering.

- K-Nearest Neighbor is a typical clustering technique

- Association rule discovery is a typical technique of clustering

- Clustering techniques do not require any parameters

- A, B

- A, B, and C

**Question 6**

0 / 3 pts

Select the correct sentence or sentences that define cosine and correlation measures.



The range of values that are possible for the cosine measure is [0, 1]



If two objects have a cosine similarity of 1, they are identical



If  $X$  and  $Y$  have a mean of 0,  $\text{corr}(X, Y) = \cos(X, Y)$



A, B



A, B and C

**Correct Answer****You Answered****Question 7**

3 / 3 pts

Select all classification techniques.

**Correct!** Decision tree**Correct!** Naïive Bayes**Correct!** Clustering Neural Networks

Decision tree, Naïve Bayes and Neural Network are commonly used classification techniques.

**Question 8****3 / 3 pts**

Review the table below.

ID	Device OS	Carrier	Device Color	Earlier Adopter (Class)
1	Android	T-Mobile	White	Y
2	Android	AT&T	Gray	Y
3	Android	AT&T	Gray	Y
4	Android	AT&T	Black	Y
5	Android	AT&T	Pink	Y
6	Android	AT&T	Pink	Y
7	IOS	AT&T	White	Y
8	IOS	AT&T	White	Y
9	IOS	AT&T	Gray	Y
10	IOS	Verizon	Black	Y
11	Android	T-Mobile	Black	N
12	Android	T-Mobile	Pink	N
13	Android	T-Mobile	Gray	N
14	Android	Verizon	Pink	N
15	IOS	Verizon	White	N
16	IOS	Verizon	White	N
17	IOS	Verizon	Gray	N
18	IOS	Verizon	Gray	N
19	IOS	Verizon	Gray	N
20	IOS	Verizon	Black	N

Based on the table, which of the following statements are correct about a decision tree model?

Correct!

- Gini index for each ID value is 0.

Correct!



Gini for Android in Device OS is the same with Gini for IOS in Device OS.

Correct!

- Gini index for the carrier is 0.1625.

Correct!

- Gini index for the overall collection of the example is 0.5.

### Question 9

3 / 3 pts

Review the table below.

ID	Device OS	Carrier	Device Color	Earlier Adopter (Class)
1	Android	T-Mobile	White	Y
2	Android	AT&T	Gray	Y
3	Android	AT&T	Gray	Y
4	Android	AT&T	Black	Y
5	Android	AT&T	Pink	Y
6	Android	AT&T	Pink	Y
7	IOS	AT&T	White	Y
8	IOS	AT&T	White	Y
9	IOS	AT&T	Gray	Y
10	IOS	Verizon	Black	Y
11	Android	T-Mobile	Black	N
12	Android	T-Mobile	Pink	N
13	Android	T-Mobile	Gray	N
14	Android	Verizon	Pink	N
15	IOS	Verizon	White	N
16	IOS	Verizon	White	N
17	IOS	Verizon	Gray	N
18	IOS	Verizon	Gray	N
19	IOS	Verizon	Gray	N
20	IOS	Verizon	Black	N

Based on the table, which of the following statements are correct about the decision tree model?

- 
- The ID is the best attribute to predict Earlier Adopter
  - The Carrier is the best attribute to predict Earlier Adopter
  - The Device OS is the best attribute to predict Earlier Adopter
  - The Device Color is the best attribute to predict Earlier Adopter

Correct!

**Question 10****3 / 3 pts**

Review the following table.

Instance	Attribute 1	Attribute 2	Attribute 3	class
1	T	T	1	Y
2	T	T	6	Y
3	T	F	5	N
4	F	F	4	Y
5	F	T	7	N
6	F	T	3	N
7	F	F	8	N
8	T	F	7	Y
9	F	T	5	N

What is the best split between Attribute 1 and Attribute 2 according to the classification rate and what is the classification error of the split attribute?

**Correct!**

- Attribute 1 is the best split, 2/9 error rate
- Attribute 1 is the best split, 4/9 error rate
- Attribute 2 is the best split, 2/9 error rate
- Attribute 2 is the best split, 4/9 error rate

**Question 11****3 / 3 pts**

ID	Device OS	Carrier	Device Color	Earlier Adopter (Class)
1	Android	T-Mobile	White	Y
2	Android	AT&T	Gray	Y
3	Android	AT&T	Gray	Y
4	Android	AT&T	Black	Y
5	Android	AT&T	Pink	Y
6	Android	AT&T	Pink	Y
7	IOS	AT&T	White	Y
8	IOS	AT&T	White	Y
9	IOS	AT&T	Gray	Y
10	IOS	Verizon	Black	Y
11	Android	T-Mobile	Black	N
12	Android	T-Mobile	Pink	N
13	Android	T-Mobile	Gray	N
14	Android	Verizon	Pink	N
15	IOS	Verizon	White	N
16	IOS	Verizon	White	N
17	IOS	Verizon	Gray	N
18	IOS	Verizon	Gray	N
19	IOS	Verizon	Gray	N
20	IOS	Verizon	Black	N

Compute the Gini index for Device Color attribute using multiway split

**Correct!**

0.491

**Correct Answers**

Between 0.49 and 0.4915

**Question 12**

3 / 3 pts

Review the following table.

Ground Truth	apple									
Predict	apple	pear	pear	apple	apple	pear	pear	pear	apple	apple

Compute F1 for 'pear' class

**Correct!** 0**Correct Answers**

0 (with margin: 0)

**Question 13**

3 / 3 pts

Review the table below.

Ground Truth	apple									
Predict	apple	pear	pear	apple	apple	pear	pear	pear	apple	apple

Select the correct precision and recall for 'apple' class.

- 
- Precision – 0.5 and recall – 0.5

**Correct!**

- Precision – 1 and recall – 0.5
- Precision – 0.5 and recall – 1
- Precision – 1 and recall – 1

**Question 14****3 / 3 pts**

Review the following table.

Instance	Attribute 1	Attribute 2	Attribute 3	class
1	T	T	1	Y
2	T	T	6	Y
3	T	F	5	N
4	F	F	4	Y
5	F	T	7	N
6	F	T	3	N
7	F	F	8	N
8	T	F	7	Y
9	F	T	5	N

Select the correct information gain pair of Attribute 1 and Attribute 2.

**Correct!**

- Attribute 1- 0.229, Attribute 2 – 0.007

- Attribute 1- 0.991, Attribute 2 – 0.007
- Attribute 1- 0.007, Attribute 2 – 0.229
- Attribute 1- 0.229, Attribute 2 – 0.991

**Question 15****3 / 3 pts**

Select all sentences that correctly define errors.

**Correct!**

- Prediction errors consist of bias, variance, and irreducible error.

**Correct!**

- Bias error results in high training error.

- Variance error results in high training error.

**Correct!**

- Irreducible error cannot be optimized.

**Question 16****3 / 3 pts**

Select Methods for good model evaluation.

**Correct!**

- Maximize values of both precision and recall

we need to maximize both precision and recall

- Maximize accuracy
- Maximize precision
- Minimize recall

Quiz Score: **44.25** out of 48

# Midterm

**Due** Oct 28 at 11:59pm      **Points** 48      **Questions** 16

**Available** Oct 26 at 12am - Oct 29 at 1am 3 days

**Time Limit** 100 Minutes

## Attempt History

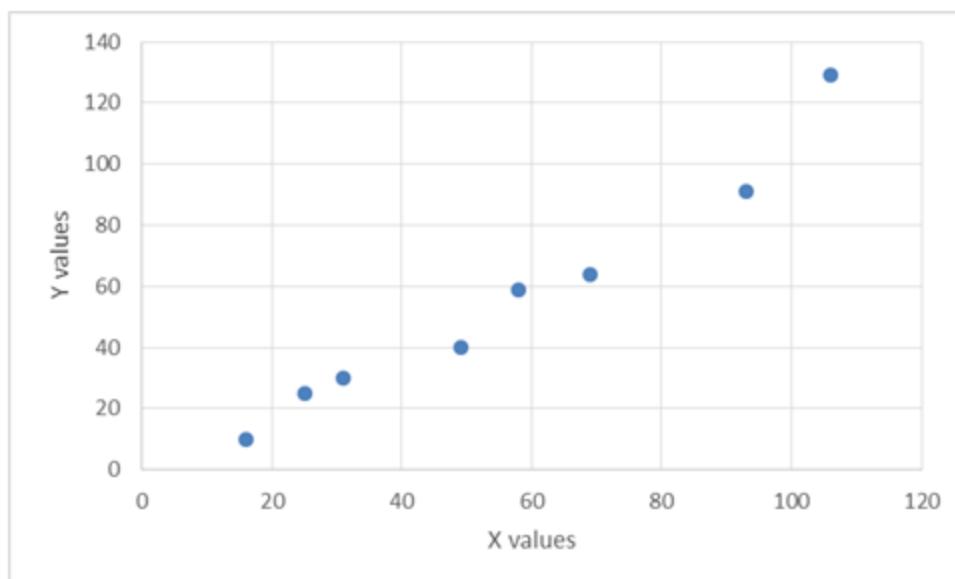
	<b>Attempt</b>	<b>Time</b>	<b>Score</b>
<b>LATEST</b>	<a href="#"><u>Attempt 1</u></a>	43 minutes	46.5 out of 48

Score for this quiz: **46.5** out of 48

Submitted Oct 26 at 9:23pm

This attempt took 43 minutes.

Question 1			3 / 3 pts
	X	Y	
A	16	10	
B	25	25	
C	31	30	
D	49	40	
E	58	59	
F	69	64	
G	93	91	
H	106	129	



Compute cosine similarity between X and Y

**Correct!**

0.991

**Correct Answers**

Between 0.95 and 1

## Question 2

1.5 / 3 pts

Select all predictive techniques.

**You Answered**

- Anomaly Detection

**Correct!**

- Regression

**Correct!**

- Classification

- Clustering

- Dimensionality Reduction

### Question 3

3 / 3 pts

Select all correct sentences which make up the clustering.

**Correct!**

- K-Nearest Neighbor is a typical clustering technique

- Association rule discovery is a typical technique of clustering

- Clustering techniques do not require any parameters

- A, B

- A, B, and C

### Question 4

3 / 3 pts

Select the correct sentence or sentences that define cosine and correlation measures.

**Correct!**

The range of values that are possible for the cosine measure is [0, 1]

If two objects have a cosine similarity of 1, they are identical

If X and Y have a mean of 0,  $\text{corr}(X,Y) = \cos(X,Y)$

A, B

A, B and C

**Correct!****Question 5****3 / 3 pts**

Select all supervised learning methodologies

Clustering

Dimensionality Reduction

Classification or categorization

Classification is a discrete data based supervised learning method.

**Correct!**

Regression

Regression is a continuous data based supervised learning method.

**Correct!****Question 6****3 / 3 pts**

Tid	Refund	Marital Status	Taxable Income	Cheat
1	Yes	Single	125K	No
2	No	Married	100K	No
3	No	Single	70K	No
4	Yes	Married	120K	No
5	No	Divorced	95K	Yes
6	No	Married	60K	No
7	Yes	Divorced	220K	No
8	No	Single	85K	Yes
9	No	Married	75K	No
10	No	Single	90K	Yes

Select all possible columns which can be set as the class columns.

Tid

**Correct!**

Refund

**Correct!**

Marital Status

**Correct!**

Taxable Income

**Correct!**

Cheat

### Question 7

3 / 3 pts

ID	Device OS	Carrier	Device Color	Earlier Adopter (Class)
1	Android	T-Mobile	White	Y
2	Android	AT&T	Gray	Y
3	Android	AT&T	Gray	Y
4	Android	AT&T	Black	Y
5	Android	AT&T	Pink	Y
6	Android	AT&T	Pink	Y
7	IOS	AT&T	White	Y
8	IOS	AT&T	White	Y
9	IOS	AT&T	Gray	Y
10	IOS	Verizon	Black	Y
11	Android	T-Mobile	Black	N
12	Android	T-Mobile	Pink	N
13	Android	T-Mobile	Gray	N
14	Android	Verizon	Pink	N
15	IOS	Verizon	White	N
16	IOS	Verizon	White	N
17	IOS	Verizon	Gray	N
18	IOS	Verizon	Gray	N
19	IOS	Verizon	Gray	N
20	IOS	Verizon	Black	N

Compute the Gini index for Device Color attribute using multiway split

**Correct!**

0.491

**Correct Answers**

Between 0.49 and 0.4915

### Question 8

**3 / 3 pts**

Select the statements below that are computed as part of the performance metric.

**Correct!**

Precision can be computed using True Positive and False Positive.

Recall can be computed using True Positive and False Positive.

**Correct!**

F1 can be computed using True Positive and False Positive, and False Negative.

Accuracy can be computed using True Positive, False Positive, and False Negative.

### Question 9

**3 / 3 pts**

Review the following table.

Instance	Attribute 1	Attribute 2	Attribute 3	class
1	T	T	1	Y
2	T	T	6	Y
3	T	F	5	N
4	F	F	4	Y
5	F	T	7	N
6	F	T	3	N
7	F	F	8	N
8	T	F	7	Y
9	F	T	5	N

When is the best split for Attribute 3 according to the information gain?

**Correct!**

The split point equal to 2.0

The split point equal to 3.5

The split point equal to 5.5

The split point equal to 6.5

The split point equal to 7.5

### Question 10

3 / 3 pts

Select all classification techniques.

Correct!

Decision tree

Correct!

Naïive Bayes

Correct!

Neural Networks

Decision tree, Naïive Bayes and Neural Network are commonly used classification techniques.

### Question 11

3 / 3 pts

Review the following table.

Instance	Attribute 1	Attribute 2	Attribute 3	class
1	T	T	1	Y
2	T	T	6	Y
3	T	F	5	N
4	F	F	4	Y
5	F	T	7	N
6	F	T	3	N
7	F	F	8	N
8	T	F	7	Y
9	F	T	5	N

Select the best split among attribute 1, attribute 2, and attribute 3, according to the information gain.

**Correct!**

Attribute 1

Attribute 2

Attribute 3

### Question 12

3 / 3 pts

Select all sentences that correctly define errors.

**Correct!**

Errors committed on training records are training errors.

**Correct!**

Generalization error are expected errors of the model on previously unseen records.

Low training errors guarantee low generalization errors.

Generalization errors can be reduced by increasing the model complexity.

**Question 13****3 / 3 pts**

True or False? The tree induction algorithm C4.5 is suitable for large dataset.

True

False

C4.5 is suitable for small data set.

**Correct!****Question 14****3 / 3 pts**

Review the following table.

Instance	Attribute 1	Attribute 2	Attribute 3	class
1	T	T	1	Y
2	T	T	6	Y
3	T	F	5	N
4	F	F	4	Y
5	F	T	7	N
6	F	T	3	N
7	F	F	8	N
8	T	F	7	Y
9	F	T	5	N

What is the best split between Attribute 1 and Attribute 2 according to the Gini index? What is the Gini index of the split attribute?

**Correct!**

Attribute 1 is the best split and Gini index is 0.3444

Attribute 1 is the best split and Gini index is 0.4889

Attribute 2 is the best split and Gini index is 0.3444

Attribute 2 is the best split and Gini index is 0.4889

### Question 15

3 / 3 pts

Review the table below.

Ground Truth	apple									
Predict	apple	pear	pear	apple	apple	pear	pear	pear	apple	apple

Select the correct precision and recall for 'apple' class.

Precision – 0.5 and recall – 0.5

Precision – 1 and recall – 0.5

Precision – 0.5 and recall – 1

Precision – 1 and recall – 1

**Correct!**

### Question 16

3 / 3 pts

Select all statements that correctly address underfitting and overfitting.

**Correct!**

Underfitting can be reduced by reducing bias error.

Overfitting can be reduced by reducing bias error.

Underfitting can be reduced by reducing variance error.

**Correct!**

Overfitting can be reduced by reducing variance error.

Quiz Score: **46.5** out of 48

# Midterm

**Due** Oct 28 at 11:59pm      **Points** 48      **Questions** 16

**Available** Oct 26 at 12am - Oct 29 at 1am 3 days

**Time Limit** 100 Minutes

## Attempt History

	<b>Attempt</b>	<b>Time</b>	<b>Score</b>
<b>LATEST</b>	<a href="#"><u>Attempt 1</u></a>	54 minutes	48 out of 48

Score for this quiz: **48** out of 48

Submitted Oct 28 at 12:03am

This attempt took 54 minutes.

### Question 1

3 / 3 pts

Select all possible use-cases of clustering

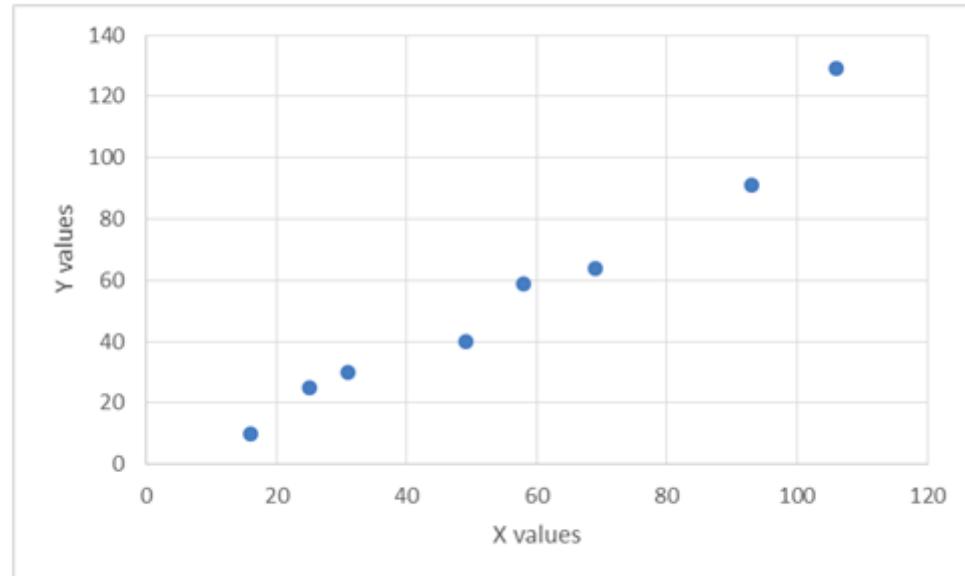
- Credit card fraud detection
- Market segmentation
- Summarize news
- A, B, and C
- A, B

Correct!

### Question 2

3 / 3 pts

	X	Y
A	16	10
B	25	25
C	31	30
D	49	40
E	58	59
F	69	64
G	93	91
H	106	129



Select the closest answer for the correlation between X and Y

0.5

-0.5

-1

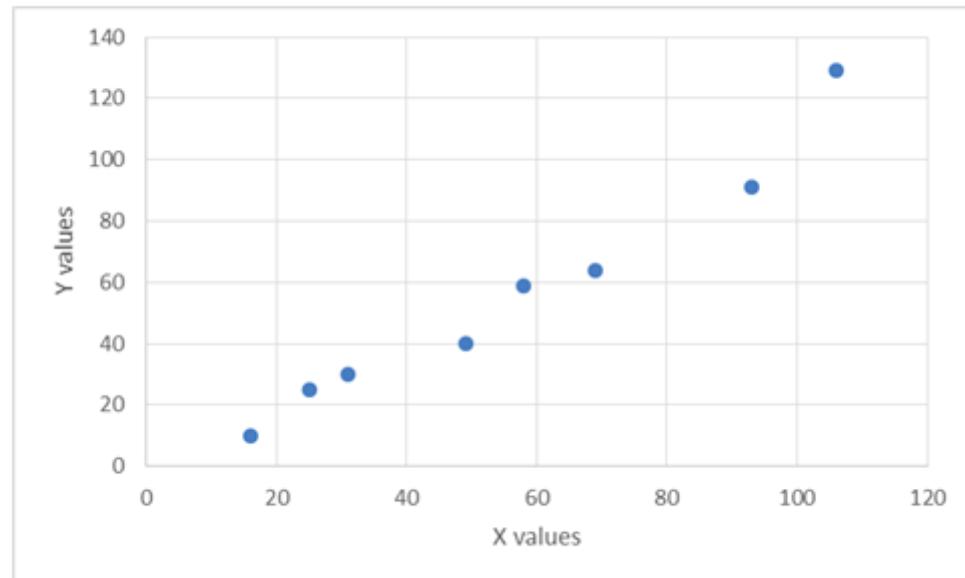
**Correct!** 1 0**Question 3****3 / 3 pts**

Select all correct sentences which make up the clustering.

- K-Nearest Neighbor is a typical clustering technique
- Association rule discovery is a typical technique of clustering
- Clustering techniques do not require any parameters
- A, B
- A, B, and C

**Question 4****3 / 3 pts**

	X	Y
A	16	10
B	25	25
C	31	30
D	49	40
E	58	59
F	69	64
G	93	91
H	106	129



Compute cosine similarity between X and Y

Correct!

0.991

Correct Answers

Between 0.95 and 1

**Question 5****3 / 3 pts**

Tid	Refund	Marital Status	Taxable Income	Cheat
1	Yes	Single	125K	No
2	No	Married	100K	No
3	No	Single	70K	No
4	Yes	Married	120K	No
5	No	Divorced	95K	Yes
6	No	Married	60K	No
7	Yes	Divorced	220K	No
8	No	Single	85K	Yes
9	No	Married	75K	No
10	No	Single	90K	Yes

Select all possible columns which can be set as the class columns.

 Tid**Correct!** Refund**Correct!** Marital Status**Correct!** Taxable Income**Correct!** Cheat**Question 6****3 / 3 pts**

Select all predictive techniques.

 Anomaly Detection**Correct!**

**Correct!** Regression**Correct!** Classification Clustering Dimensionality Reduction**Question 7****3 / 3 pts**

True or False? The tree induction algorithm C4.5 is suitable for large dataset.

 True False

C4.5 is suitable for small data set.

**Correct!****Question 8****3 / 3 pts**

Select all sentences that are correct about cross-validation.

**Correct!** Leave-one-out is a type of cross-validation. Cross-validation guarantees reduction in test error.

K-fold cross-validation requires K repetitions of the same train set on K different test sets..

**Correct!**

In cross validation the same training record can be selected multiple times.

**Question 9****3 / 3 pts**

Select the statements below that are computed as part of the performance metric.

**Correct!**

Precision can be computed using True Positive and False Positive.



Recall can be computed using True Positive and False Positive.

**Correct!**

F1 can be computed using True Positive and False Positive, and False Negative.



Accuracy can be computed using True Positive, False Positive, and False Negative.

**Question 10****3 / 3 pts**

Review the following table.

Ground Truth	apple									
Predict	apple	pear	pear	apple	apple	pear	pear	pear	apple	apple

Compute the accuracy as a fraction

**Correct!**

0.5

**orrect Answers**

0.5 (with margin: 0)

**Question 11****3 / 3 pts**

Select all statements that correctly address underfitting and overfitting.

**Correct!** Underfitting can be reduced by reducing bias error. Overfitting can be reduced by reducing bias error. Underfitting can be reduced by reducing variance error.**Correct!** Overfitting can be reduced by reducing variance error.**Question 12****3 / 3 pts**

Review the following table.

Ground Truth	apple									
Predict	apple	pear	pear	apple	apple	pear	pear	pear	apple	apple

Compute F1 for 'pear' class

**Correct!**

0

**orrect Answers**

0 (with margin: 0)

**Question 13****3 / 3 pts**

Review the following table.

Instance	Attribute 1	Attribute 2	Attribute 3	class
1	T	T	1	Y
2	T	T	6	Y
3	T	F	5	N
4	F	F	4	Y
5	F	T	7	N
6	F	T	3	N
7	F	F	8	N
8	T	F	7	Y
9	F	T	5	N

What is the best split between Attribute 1 and Attribute 2 according to the Gini index? What is the Gini index of the split attribute?

**Correct!**

- Attribute 1 is the best split and Gini index is 0.3444
- Attribute 1 is the best split and Gini index is 0.4889
- Attribute 2 is the best split and Gini index is 0.3444
- Attribute 2 is the best split and Gini index is 0.4889

**Question 14****3 / 3 pts**

Select all techniques used to measure node impurity.

**Correct!** Gini Index

Gini Index is a measure of node impurity

 average**Correct!** Entropy

Entropy is a measure of node impurity

 split**Question 15**

3 / 3 pts

A good classification model has:

 Low training error Low estimated generalization error A and B are correct

We need both low training error and low generalization error.

 None of the above**Question 16**

3 / 3 pts

Review the table below.

Ground Truth	apple									
Predict	apple	pear	pear	apple	apple	pear	pear	pear	apple	apple

Compute F1 for 'apple' class (the result should be rounded to the thousandths place)

Correct!

0.667

Correct Answers      Between 0.6 and 0.7

Quiz Score: **48** out of 48

# Module 7 Practice Quiz

**Due** No due date    **Points** 0    **Questions** 10  
**Time Limit** None    **Allowed Attempts** Unlimited

[Take the Quiz Again](#)

## Attempt History

	<b>Attempt</b>	<b>Time</b>	<b>Score</b>
LATEST	<a href="#">Attempt 1</a>	less than 1 minute	0 out of 0

Submitted Dec 8 at 7:05pm

Inanswered	<b>Question 1</b>	0 / 0 pts
	Select the main methods used for Monte carlo methods?	
Correct Answer	<input type="checkbox"/> Exploration	
	<input type="checkbox"/> Dynamic Programming	
Correct Answer	<input type="checkbox"/> Simulation	
	<input type="checkbox"/> All of the above	

Inanswered	<b>Question 2</b>	0 / 0 pts
	How can you mitigate the weakness of greedy policy for Monte Carlo Methods?	

**Correct Answer**

- By using  $\epsilon$ -greedy policy
- Greedy policy is always optimal
- None of the above
- All of the above

**Inanswered****Question 3**

0 / 0 pts

What is the weakness of greedy policy for Monte Carlo Methods?

- $V\pi$  not enough for policy improvement
- Greedy policy won't explore all actions
- They don't have full knowledge of the system
- None of the above

**Correct Answer****Question 4**

0 / 0 pts

How do you get the value function MC?

- Use both policy evaluation and policy improvement
- Need exact model of environment
- By averaging sample returns
- None of the above

**Correct Answer**

In answered

**Question 5**

0 / 0 pts

In what case can you use DP for solving RL problem?

Need to try all actions to find the optimal one

When you can divide the problem into sub problem

All of the above

None of the above

orrect Answer

In answered

**Question 6**

0 / 0 pts

Select all the methods used in Dynamic Programming in RL?

Start with an arbitrary policy

Need perfect model of environment

Repeat evaluation and improvement until convergence

None of the above

All of the above

orrect Answer

In answered

**Question 7**

0 / 0 pts

For dynamic programming what do we mean by Policy Improvement?

- Improve  $V\pi$  from  $\pi$
- Learn From Sample Episodes of Simulated Experience
- Compute  $\pi$  from  $V\pi$

**Correct Answer**

- Improve  $\pi$  from  $V\pi$
- All of the above

**Inanswered****Question 8**

0 / 0 pts

For dynamic programming, what do we mean by Policy Evaluation?

**Correct Answer**

- Compute  $V\pi$  from  $\pi$
- Learn From Sample Episodes of Simulated Experience
- Maximize cumulative reward in the long run
- Compute  $\pi$  from  $V\pi$
- None of the above

**Inanswered****Question 9**

0 / 0 pts

Select all the main Difference between RL of MDP?

- RL does not have a Reward Function  $r(s)$
- MDP starts with an arbitrary policy



MDP have Set of States (S), Set of Actions (A), Initial State S0 but RL does not

**Correct Answer**

- RL doesn't assume that you have a model
- All of the above

**Inanswered****Question 10****0 / 0 pts**

How does RL overcome not having complete model of the system?

- By learning from the environment
- Learning From Sample Episodes of Simulated Experience
- Maximizing cumulative reward in the long run

**Correct Answer**

- All of the above

# Knowledge Check (Solving Reinforcement Problems)

**Due** No due date    **Points** 0    **Questions** 1  
**Time Limit** None    **Allowed Attempts** Unlimited

[Take the Quiz Again](#)

## Attempt History

	<b>Attempt</b>	<b>Time</b>	<b>Score</b>
LATEST	<a href="#">Attempt 1</a>	less than 1 minute	0 out of 0

Submitted Dec 8 at 7:05pm

In answered	<b>Question 1</b>	0 / 0 pts
	What is the proper approach to solve reinforcement learning using Monte Carlo method?	
	<input type="radio"/> Greedy	
	<input type="radio"/> Divided-and-Conquer	
Correct Answer	<input checked="" type="radio"/> Brute Force	
	<input type="radio"/> Dynamic Programming	
	<input type="radio"/> None of the above	

# Knowledge Check (Reinforcement Learning)

**Due** No due date    **Points** 0    **Questions** 1  
**Time Limit** None    **Allowed Attempts** Unlimited

[Take the Quiz Again](#)

## Attempt History

	<b>Attempt</b>	<b>Time</b>	<b>Score</b>
LATEST	<a href="#">Attempt 1</a>	less than 1 minute	0 out of 0

Submitted Dec 8 at 7:04pm

In answered	<b>Question 1</b>	0 / 0 pts
	Select the correct statement about reinforcement learning.	
	<input type="radio"/> Reinforcement Learning is a supervised learning.	
	<input type="radio"/> Reinforcement Learning is an unsupervised learning.	
	<input type="radio"/> Reinforcement Learning is a semi-supervised learning	
Correct Answer	<input type="radio"/> None of above	

# Knowledge Check (Markov Decision)

<b>Due</b>	No due date	<b>Points</b>	0	<b>Questions</b>	1
<b>Time Limit</b>	None	<b>Allowed Attempts</b>	Unlimited		

[Take the Quiz Again](#)

## Attempt History

	<b>Attempt</b>	<b>Time</b>	<b>Score</b>
LATEST	<a href="#">Attempt 1</a>	less than 1 minute	0 out of 0

Submitted Dec 8 at 7:04pm

In answered	<b>Question 1</b>	0 / 0 pts
	What is the goal of Markov Decision Process?	
	<input type="radio"/> Minimize cumulative reward in the long run	
	<input type="radio"/> Minimize cumulative reward in the current state	
Correct Answer	<input checked="" type="radio"/> Maximize cumulative reward in the long run	
	<input type="radio"/> Maximize cumulative reward in the current state	
	<input type="radio"/> None of the above	

# Module 6 Practice Quiz

**Due** No due date    **Points** 0    **Questions** 9  
**Time Limit** None    **Allowed Attempts** Unlimited

[Take the Quiz Again](#)

## Attempt History

	<b>Attempt</b>	<b>Time</b>	<b>Score</b>
LATEST	<a href="#">Attempt 1</a>	less than 1 minute	0 out of 0

Submitted Dec 8 at 7:04pm

In answered	<b>Question 1</b>	0 / 0 pts
	Which of the following statement(s) correctly represents a real neuron?	
	<input type="radio"/> A neuron can have a single input and a single output only	
	<input type="radio"/> A neuron can have multiple inputs but a single output only	
	<input type="radio"/> A neuron can have a single input but multiple outputs	
	<input type="radio"/> A neuron can have multiple inputs and multiple outputs	
Correct Answer	<input checked="" type="radio"/> All of the above statements are valid	

In answered	<b>Question 2</b>	0 / 0 pts
	In a neural network, knowing the weight and bias of each neuron is the most important step. If you can somehow get the correct value of weight	

and bias for each neuron, you can approximate any function. What would be the best way to approach this?

- 
- Assign random values and pray to God they are correct

---

  - Search every possible combination of weights and biases till you get the best value
- 

Correct Answer

- Iteratively check that after assigning a value how big of an error you make, and slightly change the assigned values values to decrease the error

---

- None of these

Inanswered

### Question 3

0 / 0 pts

Which of the following methods is designed to best fit the data in a least-squares sense?

- 
- Fisher's linear discriminant

---

  - Multiple discriminant analysis

---

  - Principle Component Analysis (PCA)

---

  - None of the above

Correct Answer

Inanswered

### Question 4

0 / 0 pts

"Convolutional Neural Networks can perform various types of transformation (rotations or scaling) in an input."

True**Correct Answer** False**Inanswered****Question 5****0 / 0 pts**

Which of the following architecture has feedback connections?

**Correct Answer** Recurrent Neural network Convolutional Neural Network Restricted Boltzmann Machine None of the above**Inanswered****Question 6****0 / 0 pts**

Prompt:  $Y = ax^2 + bx + c$  (polynomial equation of degree 2)

Can this equation be represented by a neural network of single hidden layer with linear threshold?

**Correct Answer** Yes No**Inanswered****Question 7****0 / 0 pts**

For an image recognition problem (recognizing a cat in a photo), which architecture of neural network would be better suited to solve the problem?

---

Multi Layer Perceptron

---

Correct Answer

Convolutional Neural Network

---

Recurrent Neural network

---

Perceptron

---

Inanswered

## Question 8

0 / 0 pts

Answer location in instruction: For which application Autoencoder Neural Networks cannot be used directly?

---

Denoising

---

Data compression

---

Representing an RGB image as a GrayScale Image

---

Correct Answer

Clustering

---

Inanswered

## Question 9

0 / 0 pts

What type of learning algorithm is an Autoencoder?

---

supervised

---

**orrect Answer**

- unsupervised with backpropagation
- distributive
- unsupervised without any backpropagation
- none of the above

# Knowledge Check: Deep Learning Introduction

**Due** No due date    **Points** 0    **Questions** 4  
**Time Limit** None    **Allowed Attempts** Unlimited

[Take the Quiz Again](#)

## Attempt History

	<b>Attempt</b>	<b>Time</b>	<b>Score</b>
LATEST	<a href="#">Attempt 1</a>	less than 1 minute	0 out of 0

Submitted Dec 8 at 7:03pm

In answered	<b>Question 1</b>	0 / 0 pts
	Give the full form of LSTM in the context of deep learning	
	<input type="radio"/> Long Short Time Memory	
Correct Answer	<input checked="" type="radio"/> Long Short Term Memory	
	<input type="radio"/> Long Short Time Module	
	<input type="radio"/> Lucky Short Time Management	

In answered	<b>Question 2</b>	0 / 0 pts
	Restricted Boltzmann Machine (RBM) is a type of	

**Correct Answer**

- Belief Network
- Autoencoder
- Convolutional Neural Network
- Recurrent Neural Network

**Inanswered****Question 3**

0 / 0 pts

Neural network can be used to extract knowledge for complex data.

**Correct Answer**

- True
- False

**Inanswered****Question 4**

0 / 0 pts

Select the reason for using deep learning to recommendation systems

**Correct Answer**

- To overcome cold start problems
- Deep learning is easy to implement
- Deep learning always performs better for recommendation systems
- None of the above

# Module 5 Practice Quiz

**Due** No due date

**Points** 0

**Questions** 10

**Time Limit** None

## Attempt History

	<b>Attempt</b>	<b>Time</b>	<b>Score</b>
<b>LATEST</b>	<a href="#"><u>Attempt 1</u></a>	less than 1 minute	0 out of 0

Submitted Dec 8 at 7:03pm

In answered

### Question 1

0 / 0 pts

Which of the following is true?



If  $\{X\} \rightarrow \{Y\}$  is an association rule, then  $\{x\}$  and  $\{Y\}$  are positively correlated.



If  $\{X\} \rightarrow \{Y\}$  is an association rule, then  $\{x\}$  and  $\{Y\}$  are at least not negatively correlated.



If both  $\{X\} \rightarrow \{Y\}$  and  $\{Y\} \rightarrow \{X\}$  are association rules, then  $\{x\}$  and  $\{Y\}$  are positively correlated.



If both  $\{x\}$  and  $\{Y\}$  are correlated, then  $\{Y\} \rightarrow \{X\}$  is a strong association rule.

Correct Answer

Association does not imply correlation.

Unanswered

**Question 2**

0 / 0 pts

Example of market basket transactions.

Customer ID Transaction ID Items Bought

1	1	{a, d, e}
1	24	{a, b, c, e}
2	12	{a, b, d, e}
2	31	{a, c, d, e}
3	15	{b, c, e}
3	22	{b, d, e}
4	29	{c, d}
4	40	{a, b, c}
5	33	{a, d, e}
5	38	{a, b, e}

Compute the support for itemset {a,e} by treating each transaction ID as a market basket.

You Answered

Correct Answers

- 0.6 (with margin: 0)
- 0 (with margin: 0)
- 0 (with margin: 0)

Unanswered

**Question 3****0 / 0 pts**

Example of market basket transactions.

Customer ID Transaction ID Items Bought

1	1	{a, d, e}
1	24	{a, b, c, e}
2	12	{a, b, d, e}
2	31	{a, c, d, e}
3	15	{b, c, e}
3	22	{b, d, e}
4	29	{c, d}
4	40	{a, b, c}
5	33	{a, d, e}
5	38	{a, b, e}

Compute the confidence for the association

rules  $\{a, e\} \rightarrow \{d\}$ , by treating each transaction ID as a market basket.

Answered

**Correct Answers**

0.67 (with margin: 0)

0 (with margin: 0)

0 (with margin: 0)

**Inanswered****Question 4****0 / 0 pts**

Example of market basket transactions.

Customer ID Transaction ID Items Bought

1 1 {a, d, e}

1 24 {a, b, c, e}

2 12 {a, b, d, e}

2 31 {a, c, d, e}

3 15 {b, c, e}

3 22 {b, d, e}

4 29 {c, d}

4 40 {a, b, c}

5 33 {a, d, e}

5 38 {a, b, e}

Compute the support for itemset {a,e} by treating each customer ID as a market basket. Each item should be treated as a binary variable (1 if an item appears in at least one transaction bought by the customer, and 0 otherwise.)

**0 Answered****Correct Answers**

- 0.6 (with margin: 0)
- 0 (with margin: 0)
- 0 (with margin: 0)
- 0 (with margin: 0)

**0 unanswered****Question 5****0 / 0 pts**

Example of market basket transactions.

Customer ID Transaction ID Items Bought

1	1	{a, d, e}
1	24	{a, b, c, e}
2	12	{a, b, d, e}
2	31	{a, c, d, e}
3	15	{b, c, e}
3	22	{b, d, e}
4	29	{c, d}
4	40	{a, b, c}
5	33	{a, d, e}
5	38	{a, b, e}

Compute the confidence for the association rules  $\{a, e\} \rightarrow \{d\}$  by treating each customer ID as a market basket. Each item should be treated as a binary variable (1 if an item appears in at least one transaction bought by the customer, and 0 otherwise.)

0 Answered

Correct Answers

- 1 (with margin: 0)
- 0 (with margin: 0)
- 0 (with margin: 0)

Inanswered

## Question 6

0 / 0 pts

Association rule is symmetric, which means confidence of rules  $A \rightarrow B$  and  $B \rightarrow A$  are the same.

True

False

Correct Answer

Inanswered

## Question 7

0 / 0 pts

Association rule are not transitive because.

Confidence of the rules  $A \rightarrow B$  and  $B \rightarrow A$  are the same.

If Confidence of the rules  $A \rightarrow B$  and  $B \rightarrow C$  are larger than some threshold, minconf; so is the confidence  $A \rightarrow C$

**Correct Answer**

Given Confidence of the rules  $A \rightarrow B$  and  $B \rightarrow C$  are larger than some threshold,  $\text{minconf}$ ; it is possible that  $A \rightarrow C$  has a confidence less than  $\text{minconf}$ .

---

All of the above.

---

None of the above.

**Inanswered****Question 8**

0 / 0 pts

What is the confidence for the rules  $A \rightarrow \emptyset$

**Correct Answer**

---

100%

---

50%

---

0%

---

None of the above.

**Inanswered****Question 9**

0 / 0 pts

Assume that the largest frequent itemset is of size  $k$ .

How many passes does the apriori algorithm need in worst case?

---

$k - 1$

---

$k$

---

**Correct Answer**

$k + 1$

---

k<sup>2</sup> 2k 2<sup>k - 1</sup>

Unanswered

**Question 10**

0 / 0 pts

Assume that the largest frequent itemset is of size k.

There are at least how many frequent itemsets?

 k - 1 k k + 1 k<sup>2</sup> 2k

Correct Answer

 2<sup>k - 1</sup>

# Knowledge Check: Apriori Principle

**Due** No due date**Points** 3**Questions** 3**Time Limit** None**Allowed Attempts** Unlimited[Take the Quiz Again](#)

## Attempt History

	<b>Attempt</b>	<b>Time</b>	<b>Score</b>
LATEST	<a href="#">Attempt 1</a>	less than 1 minute	0 out of 3

Submitted Dec 8 at 7:03pm

Unanswered

**Question 1**

0 / 1 pts

Select all true statements about common strategy adopted by many association rule mining algorithms.

- Generate all itemsets whose support is greater or equal to minsup.
- Pay attention to frequent itemsets only.
- D&B are incorrect.
- None of the above.



Generate high confidence rules from each frequent itemset, where each rule is a binary partitioning of a frequent itemset.

- A&B are correct.

- Disregard strong rules.

**Correct Answer**

- A&C are correct.

**Inanswered****Question 2**

0 / 1 pts

Select all true statements about apriori algorithm.

**Correct Answer**

- All of the above.



Algorithm eliminates all candidate itemsets whose support counts are less than minsup.



Algorithm needs to make an additional pass over dataset.



Algorithm terminates when there are no new frequent itemsets generated.

**Inanswered****Question 3**

0 / 1 pts

Select all true statement about candidate generation and pruning.



Generates new candidate k-itemsets based on frequent (k-1) itemsets found in previous iteration.



A&D are correct.



Generate same candidate itemset multiple times.



Eliminates some of candidate k-itemsets using support-based pruning strategy.

**Correct Answer**

- A&C are correct.
- Generate as many candidates as possible.
- None of the above.
- B&C are correct.

# Knowledge Check: Basic Concepts

<b>Due</b>	No due date	<b>Points</b>	1	<b>Questions</b>	2
<b>Time Limit</b>	None	<b>Allowed Attempts</b>	Unlimited		

[Take the Quiz Again](#)

## Attempt History

	<b>Attempt</b>	<b>Time</b>	<b>Score</b>
LATEST	<a href="#">Attempt 1</a>	less than 1 minute	0 out of 1

Submitted Dec 8 at 7:02pm

Inanswered	<b>Question 1</b>	0 / 1 pts
	Why do we need to use confidence and support together?	
	<input type="radio"/> Support alone doesn't give us any information.	
Correct Answer	<input checked="" type="radio"/> Support may occur simply by chance.	
	<input type="radio"/> Confidence is the most important.	
	<input type="radio"/> None of the above.	

Inanswered	<b>Question 2</b>	0 / 0 pts
	Select the type of rules that are interesting.	
	<input type="radio"/> A rule that has reasonably high support but low confidence	

**Correct Answer**

- A rule that has low support and low confidence
- A rule that has low support and high confidence
- All of the above

# Module 4 Practice Quiz

**Due** No due date

**Points** 10

**Questions** 10

**Time Limit** None

**Allowed Attempts** Unlimited

## Instructions

10 questions!

[Take the Quiz Again](#)

Submitted Dec 8 at 7:02pm

Inanswered

### Question 1

0 / 1 pts

Select all clustering algorithms.

Correct Answer

k-Mean

k-NN

Correct Answer

DBSCAN

Correct Answer

Hierarchical

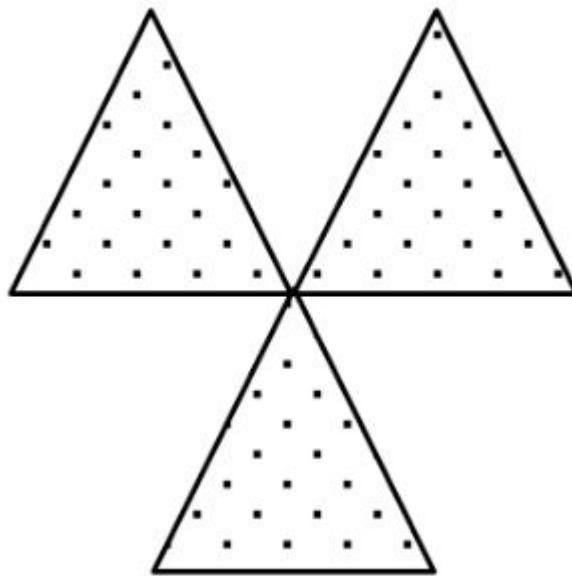
None of above

Inanswered

### Question 2

0 / 1 pts

Identify the number of clusters using the center-based, contiguity-based, and density-based definitions for below figure.



center-based: 1 cluster, contiguity-based: 2 clusters, density-based: 2 cluster



center-based: 2 cluster, contiguity-based: 2 clusters, density-based: 1 cluster

**Correct Answer**

center-based: 3 cluster, contiguity-based: 1 clusters, density-based: 3 cluster



center-based: 2 cluster, contiguity-based: 2 clusters, density-based: 2 cluster



None of above

Unanswered

**Question 3**

0 / 1 pts

Review the following table.

	A	B	C	D	E
A	0	9	3	6	11
B	9	0	7	5	10
C	3	7	0	9	2
D	6	5	9	0	8
E	11	10	2	8	0

Select the correct single-linkage (Min) hierarchical clustering result (the number of clusters is 2) using the distance matrix.

Correct Answer

- cluster1: {A, C, E} and cluster2: {B, D}
- cluster1: {A, B, D} and cluster2: {C, E}
- cluster1: {A, B, C} and cluster2: {D, E}
- cluster1: {A, B, E} and cluster2: {C, D}
- None of above

Unanswered

**Question 4**

0 / 1 pts

Review the following table.

Select the correct complete-linkage (Max) hierarchical clustering result (the number of clusters is 3).

- cluster1: {A, C, E}, cluster2: {B}, and cluster3: {D}
- cluster1: {A}, cluster2: {B, D}, and cluster3: {C, E}
- cluster1: {A, B}, cluster2: {C}, and cluster3: {D, E}
- cluster1: {A, B, E} and cluster2: {C}, and cluster3: {D}
- None of above

Correct Answer

Inanswered

### Question 5

0 / 1 pts

Review the following table.

Select the correct DBSCAN clustering result with the parameters below.

Eps: 3

Minimum samples (MinPts):2

Distance measurement: Euclidean distance

- 
- cluster1: {A, C}, cluster2: {B} cluster3: {E}, and cluster4: {D}
- 

Correct Answer

- cluster1: {A, C E}, B and D are noise points
  - cluster1: {A, B}, cluster2: {C}, and cluster3: {D}, cluster4: {E}
  - cluster1: {A}, cluster2: {B, E}, cluster3: {C}, and cluster4: {D}
  - None of above
- 

Inanswered

## Question 6

0 / 1 pts

Review the following table.

How many core samples are there if we use DBSCAN clustering with the parameters below.

Eps: 3

Minimum samples (MinPts):2

Not Answered

Correct Answers

5 (with margin: 0)

Inanswered

## Question 7

0 / 1 pts

Review the following table.

Select the correct DBSCAN clustering result with the parameters below.

Eps: 3

Minimum samples (MinPts): 2

Distance measurement: Euclidean distance

- cluster1: {A,B,C}, cluster2: {D,E,F}
- cluster1: {A,B}, cluster2: {C,D,E}, cluster3: {F}
- cluster1: {A,B,C}, cluster2: {D, E}
- cluster1: {A}, cluster2: {B}, cluster3: {C}, cluster4: {D}, cluster5: {E}
- None of above

Correct Answer

In answered

### Question 8

0 / 1 pts

Review the following table.

Compute the average of the silhouette score (ignore the noise) to evaluate the DBSCAN clustering result with the parameters below.

Eps: 3

Minimum sample:2

Distance measurement: Euclidean distance

In Answered

Correct Answers

Between 0.86 and 0.87

In answered

### Question 9

0 / 1 pts

Select all correct sentences that are true regarding k-Means clustering.

Correct Answer

The shape of each cluster is a circle

Robust to outliers after normalization processing

Correct Answer

Bisecting k-Means can solve the initial centroids issue

k-Means is not NP problem because time complexity is  $O(n*k*l*d)$ , where n is the number of points, k is the number of clusters, l is number of iterations, and d is number of attributes

None of above

In answered

### Question 10

0 / 1 pts

Select all correct sentences about DBSCAN.

DBSCAN works well in case of high dimensional data

**Correct Answer**

DBSCAN works well in geometrical shape data

**Correct Answer**

DBSCAN can handle clusters of arbitrary sizes

**Correct Answer**

DBSCAN is able to identify noise while clustering

None of above

# Knowledge Check: DBSCAN

Due	No due date	Points	2	Questions	2
Time Limit	None	Allowed Attempts	Unlimited		

## Instructions

2 questions!

[Take the Quiz Again](#)

## Attempt History

	Attempt	Time	Score
LATEST	<a href="#">Attempt 1</a>	less than 1 minute	0 out of 2

Submitted Dec 8 at 7:02pm

In answered	<b>Question 1</b>	0 / 1 pts
<p>What type of clustering is DBSCAN?</p> <p><input type="radio"/> Center-based</p> <p><input type="radio"/> Contiguous</p> <p><input type="radio"/> Well-separated</p> <p><input checked="" type="radio"/> None of above</p>		
Correct Answer		

In answered	<b>Question 2</b>	0 / 1 pts
<p></p>		

Select all that are not true about DBSCAN.

it has trouble when the clusters have widely varying densities

it is very sensitive to noise and to the size of the data

DBSCAN has a complexity of order  $O(n^3)$

it can be expensive when the computation of nearest neighbors requires computing all pairwise proximities

A and B are correct

B and C are correct

A, C and D are correct

**Correct Answer**

# Knowledge Check: Cluster Validity

<b>Due</b>	No due date	<b>Points</b>	2	<b>Questions</b>	2
<b>Time Limit</b>	None	<b>Allowed Attempts</b>	Unlimited		

## Instructions

2 questions!

[Take the Quiz Again](#)

## Attempt History

	<b>Attempt</b>	<b>Time</b>	<b>Score</b>
LATEST	<a href="#">Attempt 1</a>	less than 1 minute	0 out of 2

Submitted Dec 8 at 7:01pm

In answered	<b>Question 1</b>	0 / 1 pts
	What is the purpose of cluster analysis?	
	<input type="radio"/> To avoid finding patterns in noise	
	<input type="radio"/> To compare clustering algorithms	
	<input type="radio"/> To compare two sets of clusters	
Correct Answer	<input checked="" type="radio"/> all of the above	
	<input type="radio"/> None of the above	

that's one of the reasons for cluster validation

Inanswered

## Question 2

0 / 1 pts

Select different aspects of cluster validation.

Correct Answer

Comparing the results of a cluster analysis to externally known results.

To improve the complexity of the algorithm

Correct Answer

Determining the 'correct' number of clusters

None of the above

# Knowledge Check: K Means Clustering

---

<b>Due</b> No due date	<b>Points</b> 3	<b>Questions</b> 3
<b>Time Limit</b> None	<b>Allowed Attempts</b> Unlimited	

---

## Instructions

3 questions!

[Take the Quiz Again](#)

## Attempt History

	<b>Attempt</b>	<b>Time</b>	<b>Score</b>
<b>LATEST</b>	<a href="#"><u>Attempt 1</u></a>	less than 1 minute	0 out of 3

---

Submitted Dec 8 at 7:01pm

In answered	<b>Question 1</b>	0 / 1 pts
Select qualities of clusters produced by a good clustering algorithm.		
<input type="radio"/> Intra-cluster distances are minimized		
<input type="radio"/> Inter-cluster distances are maximized		
<input type="radio"/> Lower number of clusters are produced		
Correct Answer	<input checked="" type="radio"/> A and B are correct	

In answered	<b>Question 2</b>	0 / 1 pts
-------------	-------------------	-----------

Select all different type of clusters.

Well-separated

Center-based

Contiguous

Density-based

All of the above

that's one of clustering methods

Correct Answer

Inanswered

### Question 3

0 / 1 pts

Is it possible that Assignment of observations to clusters does not change between successive iterations in K-Means?

Correct Answer

Yes

No

Can't say

None of these

# Module 3 Practice Quiz

**Due** No due date    **Points** 0    **Questions** 10  
**Time Limit** None    **Allowed Attempts** Unlimited

## Instructions

10 questions!

[Take the Quiz Again](#)

## Attempt History

	<b>Attempt</b>	<b>Time</b>	<b>Score</b>
LATEST	<a href="#">Attempt 1</a>	less than 1 minute	0 out of 0

Submitted Dec 8 at 7pm

In answered

Question 1	0 / 0 pts
Which of the following statements is true for Nearest Neighbor classification?	
<input type="checkbox"/> Although Nearest Neighbor classification has the data dimensionality issue, it can be solved by the scaling technique	
<input type="checkbox"/> k-NN classifier is a typical lazy learner because it spend so much time building the model even if the 'k' is very small	
<input type="checkbox"/> Drawing a circle is a mandatory task to find neighbor classes regardless of the number of 'K'	

**Correct Answer**

Determining the optimal value of 'K' is important for k-NN classifier performance.

 None of above**Inanswered****Question 2**

0 / 0 pts

Which of the following sentences are correct for the k-Nearest Neighbor classifier?

**Correct Answer** k-NN is a memory intensive and computationally expensive algorithm**Correct Answer** k-NN is sensitive to scale of data**Correct Answer** k-NN can be applied to both classification and regression**Correct Answer** k-NN does not work well for rare events None of above**Inanswered****Question 3**

0 / 0 pts

Review the following table:

Weather	Rain	
	Yes	No
Sunny	2	8
Cloudy	5	5

Compute the probability of raining when the weather is sunny.

You Answered

Correct Answers

0.2 (with margin: 0)

Inanswered

### Question 4

0 / 0 pts

Suppose the fraction of high school students who can drive is 15% and the fraction of college students who can drive is 23%. If one-fifth of the students are college students and the rest are high school students, what is the probability that a student who can drive is a college student? Select the correct probability.

0.23

0.2

0.8

0.18

Correct Answer

0.28

Inanswered

### Question 5

0 / 0 pts

Which of the following statements are true?

Correct Answer

Perceptron is a linear classifier



XOR cannot be solved by ANN because perceptron is a linear classifier.



Sigmoid, Tanh, and Linear functions are typical activation functions for ANN.



Larger networks have less probability to get stuck in a local minimum as compared with smaller networks.



None of above

Unanswered

**Question 6**

0 / 0 pts

Which of the following are true about neural networks?

High learning rate always provides quick convergence for training

Low learning rate always provides quick convergence for training

Adaptive learning rate always provides quick convergence for training

None of above

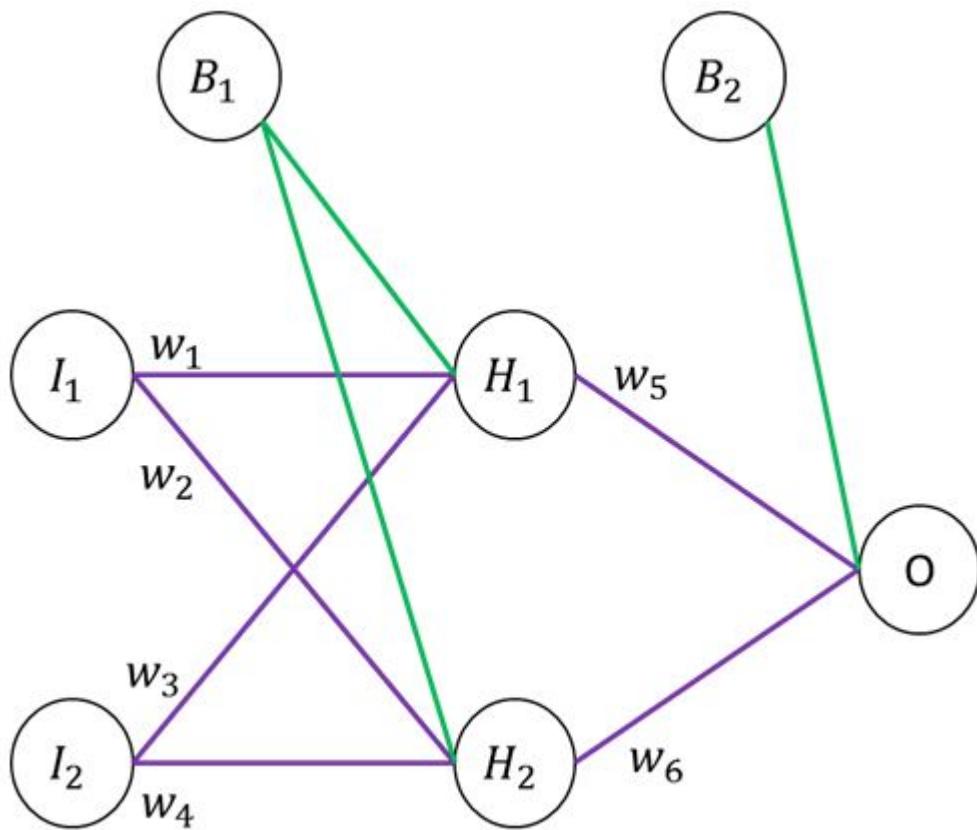
Correct Answer

Unanswered

**Question 7**

0 / 0 pts

Compute the forward pass for predicted output for O with parameters shown below. Take  $I_1 = 0.05$  and  $I_2 = 0.1$



$w_1: 0.15$

$w_2: 0.2$

$w_3: 0.25$

$w_4: 0.30$

$w_5: 0.40$

$w_6: 0.45$

$B_1: 0.35$

$B_2: 0.6$

Activation function: logistic ( $\frac{1}{1+e^{-h_i}}$ )

You Answered

Correct Answers

Between 0.75 and 0.752

Inanswered

**Question 8**

0 / 0 pts

Select the correct equation for the hyper-plane that separates the data

below using a linear SVM.

X-Axis	Y-Axis	Class
0	0	-
1	1	-
2	0	+

Correct Answer

Y=X-1

Y=X+1

Y=-X+1

Y=-X-1

None of above

Inanswered

### Question 9

0 / 0 pts

Review the SVM equation below. Select all correct sentences that are true about SVM.

$$\min_{w,b,\xi} f(w) = \frac{\|w\|^2}{2} + C \left( \sum_{i=1}^N \xi_i \right)^k$$

$$w \cdot x_i + b \geq 1 - \xi_i \quad \text{if } y_i = 1 \quad w \cdot x_i + b \leq -1 + \xi_i \quad \text{if } y_i = -1$$

Theoretically, this SVM can fall into a local minima.

Correct Answer

Theoretically, this SVM cannot fall into a local minima.

Correct Answer

Overfitting can be controlled by varying the value of parameter C.

Correct Answer

Overfitting can be controlled by varying the value of the slack variable.

- 
- None of above

Unanswered

**Question 10**

0 / 0 pts

Which of the following are linear classifiers?

- 
- Multiple Layer ANN with RELU activation.
- 
- 
- SVM with kernel
- 
- 
- Gaussian Naïve Bayes Classifier
- 
- 
- K-Nearest Neighbors

Correct Answer

- 
- None of above

# Knowledge Check: Support Vector Machines

**Due** No due date    **Points** 2    **Questions** 2  
**Time Limit** None    **Allowed Attempts** Unlimited

## Instructions

2 questions!

[Take the Quiz Again](#)

## Attempt History

	<b>Attempt</b>	<b>Time</b>	<b>Score</b>
LATEST	<a href="#"><u>Attempt 1</u></a>	less than 1 minute	0 out of 2

Submitted Dec 8 at 7pm

Inanswered **Question 1** 0 / 1 pts

The performance of SVM suffers when the dimensional of data is significantly increased.

Correct Answer  True  False

Inanswered **Question 2** 0 / 1 pts

Select the reasons for preferring decision boundaries with large margins.

**orrect Answer**

Decision boundaries with large margin tend to have better generalization errors

---



Classifiers with large margin decision boundaries are susceptible to overfitting

---

**orrect Answer**

If the margin is small, then any slight perturbation to the decision boundary can have quite a significant impact on its classification.

---

**orrect Answer**

Large margin implies there is a high confidence in classification

Large margins prevent overfitting.

# Knowledge Check: Alternative Data Mining Techniques

**Due** No due date    **Points** 3    **Questions** 3  
**Time Limit** None    **Allowed Attempts** Unlimited

## Instructions

3 questions!

[Take the Quiz Again](#)

## Attempt History

	<b>Attempt</b>	<b>Time</b>	<b>Score</b>
LATEST	<a href="#">Attempt 1</a>	less than 1 minute	0 out of 3

Submitted Dec 8 at 6:58pm

Inanswered **Question 1** 0 / 1 pts

Which of the following is true for K-NN classifiers.

Correct Answer  It does not build models explicitly

The decision of K-NN classifier is the same as that of a decision tree.

Correct Answer  Classifying unknown records is computationally expensive

All of the above

Inanswered **Question 2** 0 / 1 pts

True or False? A naïve Bayes classifier assumes that the attributes are conditionally independent given the class.

Correct Answer

True

False

Inanswered

### Question 3

0 / 1 pts

Given Bayes' Theorem  $P(A|B) = P(A,B)/P(B)$ , which of the following events can we apply Naive Bayesian Classifier to?

When the events are conditionally dependent.

When the events are conditionally independent.

Correct Answer

When  $P(B)$  is greater than zero.

# Module 2 Practice Quiz

**Due** No due date    **Points** 10    **Questions** 10  
**Time Limit** None    **Allowed Attempts** Unlimited

## Instructions

10 questions!

[Take the Quiz Again](#)

## Attempt History

	<b>Attempt</b>	<b>Time</b>	<b>Score</b>
LATEST	<a href="#">Attempt 1</a>	less than 1 minute	0 out of 10

Submitted Dec 8 at 6:57pm

In answered	<b>Question 1</b>	0 / 1 pts
Review the table below.		

ID	Device OS	Carrier	Device Color	Earlier Adopter (Class)
1	Android	T-Mobile	White	Y
2	Android	AT&T	Gray	Y
3	Android	AT&T	Gray	Y
4	Android	AT&T	Black	Y
5	Android	AT&T	Pink	Y
6	Android	AT&T	Pink	Y
7	IOS	AT&T	White	Y
8	IOS	AT&T	White	Y
9	IOS	AT&T	Gray	Y
10	IOS	Verizon	Black	Y
11	Android	T-Mobile	Black	N
12	Android	T-Mobile	Pink	N
13	Android	T-Mobile	Gray	N
14	Android	Verizon	Pink	N
15	IOS	Verizon	White	N
16	IOS	Verizon	White	N
17	IOS	Verizon	Gray	N
18	IOS	Verizon	Gray	N
19	IOS	Verizon	Gray	N
20	IOS	Verizon	Black	N

Based on the table, which of the following builds the decision tree model?

Correct Answer

- Gini index for each ID value is 0.

Correct Answer

- Gini for Android in Device OS is the same with Gini for IOS in Device OS.

Correct Answer

- Gini index for the carrier is 0.1625.

Correct Answer

- Gini index for the overall collection of the example is 0.5.

Inanswered

Question 2

0 / 1 pts

Review the table below.

ID	Device OS	Carrier	Device Color	Earlier Adopter (Class)
1	Android	T-Mobile	White	Y
2	Android	AT&T	Gray	Y
3	Android	AT&T	Gray	Y
4	Android	AT&T	Black	Y
5	Android	AT&T	Pink	Y
6	Android	AT&T	Pink	Y
7	IOS	AT&T	White	Y
8	IOS	AT&T	White	Y
9	IOS	AT&T	Gray	Y
10	IOS	Verizon	Black	Y
11	Android	T-Mobile	Black	N
12	Android	T-Mobile	Pink	N
13	Android	T-Mobile	Gray	N
14	Android	Verizon	Pink	N
15	IOS	Verizon	White	N
16	IOS	Verizon	White	N
17	IOS	Verizon	Gray	N
18	IOS	Verizon	Gray	N
19	IOS	Verizon	Gray	N
20	IOS	Verizon	Black	N

Based on the table, which of the following builds the decision tree model?

Correct Answer

- The ID is the best attribute to predict Earlier Adopter
- The Carrier is the best attribute to predict Earlier Adopter
- The Device OS is the best attribute to predict Earlier Adopter
- The Device Color is the best attribute to predict Earlier Adopter

In answered

Question 3

0 / 1 pts

ID	Device OS	Carrier	Device Color	Earlier Adopter (Class)
1	Android	T-Mobile	White	Y
2	Android	AT&T	Gray	Y
3	Android	AT&T	Gray	Y
4	Android	AT&T	Black	Y
5	Android	AT&T	Pink	Y
6	Android	AT&T	Pink	Y
7	IOS	AT&T	White	Y
8	IOS	AT&T	White	Y
9	IOS	AT&T	Gray	Y
10	IOS	Verizon	Black	Y
11	Android	T-Mobile	Black	N
12	Android	T-Mobile	Pink	N
13	Android	T-Mobile	Gray	N
14	Android	Verizon	Pink	N
15	IOS	Verizon	White	N
16	IOS	Verizon	White	N
17	IOS	Verizon	Gray	N
18	IOS	Verizon	Gray	N
19	IOS	Verizon	Gray	N
20	IOS	Verizon	Black	N

Compute the Gini index for Device Color attribute using multiway split

You Answered

Correct Answers

Between 0.49 and 0.4915

In answered

#### Question 4

0 / 1 pts

Select all sentences that correctly define errors.

Correct Answer

- Errors committed on training records are training errors.

**Correct Answer**

Generalization error are expected errors of the model on previously unseen records.

Low training errors guarantee low generalization errors.

Generalization errors can be reduced by increasing the model complexity.

**Inanswered****Question 5**

0 / 1 pts

Select all sentences that correctly define errors.

**Correct Answer**

Prediction errors consist of bias, variance, and irreducible error.

**Correct Answer**

Errors due to bias are primarily caused by models that underfit.

**Correct Answer**

Errors due to variance are caused by models that overfit.

Irreducible errors are caused by complex models.

**Inanswered****Question 6**

0 / 1 pts

Review the following table.

Ground Truth	apple									
Predict	apple	pear	pear	apple	apple	pear	pear	pear	apple	apple

Compute F1 for 'pear' class

**You Answered****Correct Answers**

0 (with margin: 0)

**Inanswered****Question 7**

0 / 1 pts

Review the following table.

Ground Truth	apple									
Predict	apple	pear	pear	apple	apple	pear	pear	pear	apple	apple

Compute the accuracy

**You Answered****Correct Answers**

0.5 (with margin: 0)

**Inanswered****Question 8**

0 / 1 pts

Review the following table.

i	ii	iii	# of instances	
			+	-
T	T	T	5	0
F	T	T	0	20
T	F	T	20	0
F	F	T	0	5
T	T	F	0	0
F	T	F	25	0
T	F	F	0	0
F	F	F	0	25

According to the classification error rate, which attribute would be chosen as the first splitting attribute?

Correct Answer

'i'

'ii'

'iii'

anyone

Inanswered

### Question 9

0 / 1 pts

Review the following table.

i	ii	iii	# of instances	
			+	-
T	T	T	5	0
F	T	T	0	20
T	F	T	20	0
F	F	T	0	5
T	T	F	0	0
F	T	F	25	0
T	F	F	0	0
F	F	F	0	25

After the first splitting and 'i' = F child node, which attribute would be chosen as the second splitting attribute?

'i'

'ii'

'iii'

anyone

Correct Answer

0 / 1 pts

In answered

### Question 10

Review the following table.

I	II	III	# of instances	
			+	-
T	T	T	5	0
F	T	T	0	20
T	F	T	20	0
F	F	T	0	5
T	T	F	0	0
F	T	F	25	0
T	F	F	0	0
F	F	F	0	25

Compute the accuracy by the resulting decision tree.

You Answered

Correct Answers

0.2 (with margin: 0)

20 (with margin: 0)

# Knowledge Check: Classification Issues

---

<b>Due</b> No due date	<b>Points</b> 4	<b>Questions</b> 4
<b>Time Limit</b> None	<b>Allowed Attempts</b> Unlimited	

---

## Instructions

4 questions!

[Take the Quiz Again](#)

## Attempt History

	<b>Attempt</b>	<b>Time</b>	<b>Score</b>
<b>LATEST</b>	<a href="#"><u>Attempt 1</u></a>	less than 1 minute	2 out of 4

---

Submitted Dec 8 at 6:57pm

### Question 1

1 / 1 pts

A good classification model has:

- A) Low training error
- B) Low generalization error
- C) A and B are correct

We need both low training error and low generalization error.

- D) None of the above

Correct!

**Question 2****1 / 1 pts**

Choose the best method for good model evaluation.

**Correct!**

- Maximize values of both precision and recall

we need to maximize both precision and recall

- Maximize accuracy

- Maximize precision

- Minimize recall

**Question 3****0 / 1 pts**

Select all that are true about estimating generalization errors.

**DU Answered**

- Generalization error is always minimized by minimizing training error

Sometime reducing training error result in an increase of generalization error.

**Correct Answer**

- 

Given two models of similar generalization errors, one should prefer the simpler model over the more complex model

**Correct Answer**

- 

Complex models can accidentally fit errors in data

- All of the above

**Question 4****0 / 1 pts**

Which of the following are true regarding bias and variance tradeoff?

**You Answered**

- Underfitting is caused by low variance alone.

We need to consider also bias when evaluating the model.

- Simple model tend to have High variance and low bias.

- Complex models overfit because of low variance

- None of the above.

**Correct Answer**

# Knowledge Check: Introduction to Classification Tasks

**Due** No due date    **Points** 4    **Questions** 4  
**Time Limit** None    **Allowed Attempts** Unlimited

## Instructions

4 questions!

[Take the Quiz Again](#)

## Attempt History

	<b>Attempt</b>	<b>Time</b>	<b>Score</b>
<b>KEPT</b>	<a href="#">Attempt 2</a>	less than 1 minute	4 out of 4
<b>LATEST</b>	<a href="#">Attempt 2</a>	less than 1 minute	4 out of 4
	<a href="#">Attempt 1</a>	1 minute	3 out of 4

Submitted Oct 27 at 7:47pm

### Question 1

1 / 1 pts

Select all techniques used to measure node impurity.

Correct!

Gini Index

Gini Index is a measure of node impurity

average

Correct!

Entropy

Entropy is a measure of node impurity

split

## Question 2

1 / 1 pts

True or False? The tree induction algorithm C4.5 is suitable for large dataset.

True

False

C4.5 is suitable for smaller data set.

Correct!

## Question 3

1 / 1 pts

Which are not issues of decision trees?

Specifying the attribute test condition.

Determining the best split.

Determine when to stop splitting.

Splitting on discrete attributes.

Discrete attributes are easy to split

Correct!

All of the above

**Question 4****1 / 1 pts**

Select all that is true about best split of node of a decision tree.

**Correct!**

- Non-homogeneous nodes have high degree of impurity

homogeneous of nodes determine its level of impurity

**Correct!**

- Nodes with homogeneous class distribution are preferred

Homogeneous nodes have low degree of impurity.

**Correct!**

- Homogeneous nodes have low degree of impurity.

Nodes with homogeneous class distribution are preferred

- None of the above

# Knowledge Check: Introduction to Classification

**Due** No due date    **Points** 2    **Questions** 2  
**Time Limit** None    **Allowed Attempts** Unlimited

## Instructions

2 questions!

[Take the Quiz Again](#)

## Attempt History

	<b>Attempt</b>	<b>Time</b>	<b>Score</b>
<b>KEPT</b>	<a href="#"><u>Attempt 2</u></a>	less than 1 minute	2 out of 2
<b>LATEST</b>	<a href="#"><u>Attempt 2</u></a>	less than 1 minute	2 out of 2
	<a href="#"><u>Attempt 1</u></a>	less than 1 minute	1 out of 2

Submitted Oct 25 at 11pm

### Question 1

1 / 1 pts

True or False? A test set is used to find a model for class attribute as a function of the values of other attributes

True

False

Correct. A test set is used to determine the accuracy of the model.

Correct!

**Question 2****1 / 1 pts**

Select all classification techniques.

**Correct!** Decision tree**Correct!** Rule-based systems**Correct!** Clustering Neural Networks

Decision tree, Rule-based systems, and Neural Network are commonly used classification techniques.

# Module 1 Practice Quiz

**Due** No due date    **Points** 10    **Questions** 10  
**Time Limit** None    **Allowed Attempts** Unlimited

## Instructions

10 questions!

LO:

- 1.1 Explain the history and purpose of data mining across multiple disciplines
- 1.2 Differentiate what is and what is not data mining
- 1.3 Describe different data mining tasks
- 1.4 Recognize attributes of data needed for data mining
- 1.5 Review and summarize data exploration techniques for use in initial data analysis

[Take the Quiz Again](#)

## Attempt History

	Attempt	Time	Score
KEPT	<a href="#">Attempt 4</a>	less than 1 minute	10 out of 10
LATEST	<a href="#">Attempt 4</a>	less than 1 minute	10 out of 10
	<a href="#">Attempt 3</a>	1 minute	8 out of 10
	<a href="#">Attempt 2</a>	15 minutes	6.25 out of 10
	<a href="#">Attempt 1</a>	4 minutes	7 out of 10

Submitted Oct 25 at 3:27am

Question 1

1 / 1 pts

Select all directly relevant fields to data mining that have influenced the breakthroughs in the field of data mining.

Correct!

- Machine Learning / Pattern Recognition

Correct!

- Statistics / Artificial Intelligence

Correct!

- Database System

- Computer Network

Correct!

## Question 2

1 / 1 pts

What factor or factors have influenced the rise of the use of data mining breakthroughs?



Services like Youtube become popular and allowed for the collection of large amounts of data.



Sensor data collection with pervasive manner became more common



Hardware became cheaper and more powerful



A, B



A, B and C

Correct!

All of them have influenced the rise of data mining.

### Question 3

1 / 1 pts

Select all correct pairs of data attributes type and the appropriate example.

Correct!

- Nominal – hair colors

Correct!

- Interval – centigrade scale

Correct!

- Ratio – Degree Kelvin or absolute

Correct!

- Ordinal – order of finishing a race

### Question 4

1 / 1 pts

Select all correct sentences that make up the classification

Correct!

- Classification is an unsupervised learning method

Correct!

- Feature extraction is a mandatory preprocessing

- Classification is a supervised learning method

Correct!

- Spam filtering is a typical classification application

- Dimensionality reduction is a typical classification technique

**Question 5****1 / 1 pts**

Select all possible use-cases of clustering

- Credit card fraud detection
- Market segmentation
- Summarize news

**Correct!**

- A, B, and C
- A, B

**Question 6****1 / 1 pts**

Select all sentences that are true.

- K-Nearest Neighbor is a typical clustering technique
- 
- Association rule discovery is a typical technique for supervised learning
- Clustering techniques do not require any parameters
- A, B

**Correct!**

- None of the above

**Question 7****1 / 1 pts**

Select all predictive techniques.

Correct!

Regression

Correct!

Classification

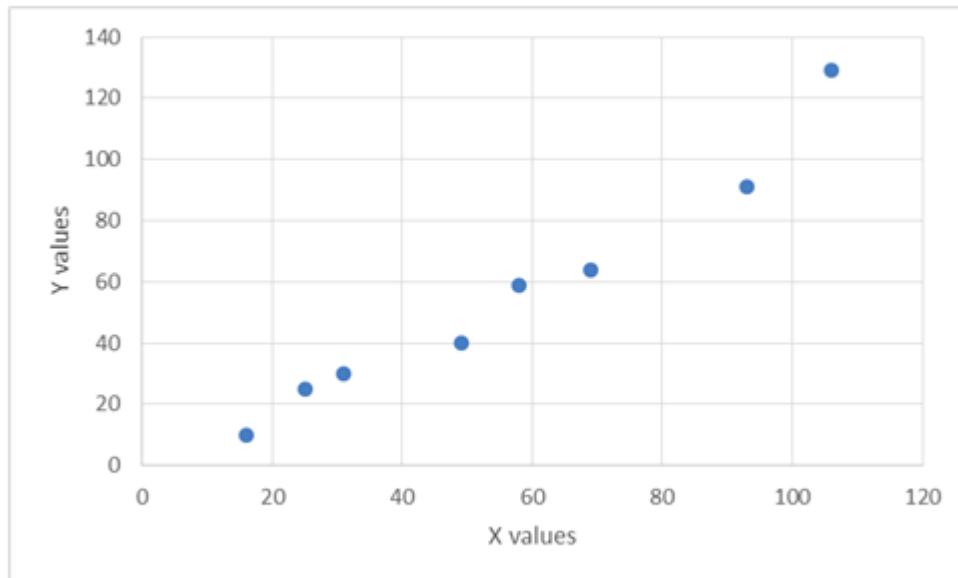
Clustering

Dimensionality Reduction

### Question 8

1 / 1 pts

	X	Y
A	16	10
B	25	25
C	31	30
D	49	40
E	58	59
F	69	64
G	93	91
H	106	129



Select the correct answer of Euclidean distance between (A and B), (D and E), and (G and H)

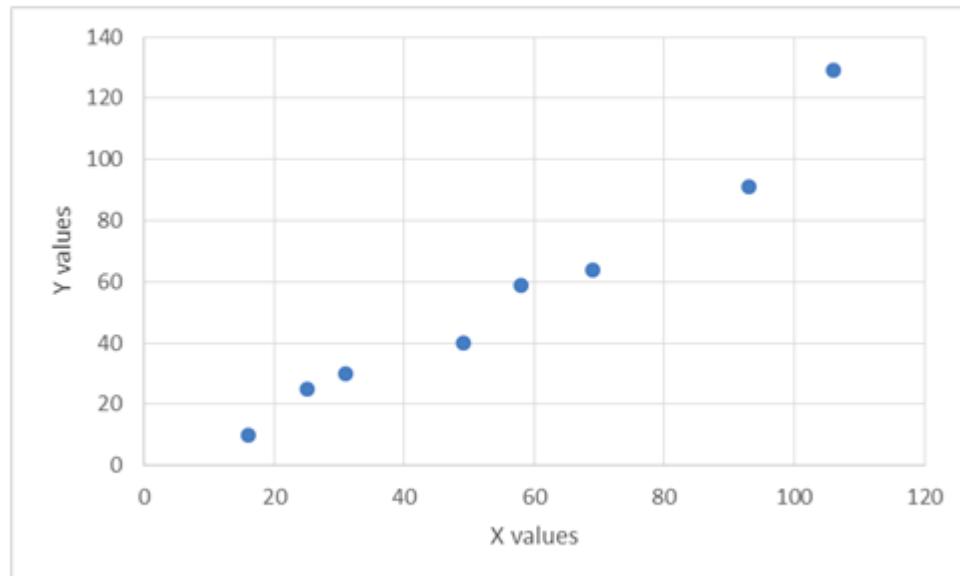
- 7.81(A and B), 39.62(D and E), 50.99(G and H)
- 17.49(A and B), 21.02(D and E), 50.99(G and H)
- 24(A and B), 28(D and E), 51(G and H)
- None

Correct!

### Question 9

1 / 1 pts

	X	Y
A	16	10
B	25	25
C	31	30
D	49	40
E	58	59
F	69	64
G	93	91
H	106	129



Compute cosine similarity between X and Y

Correct!

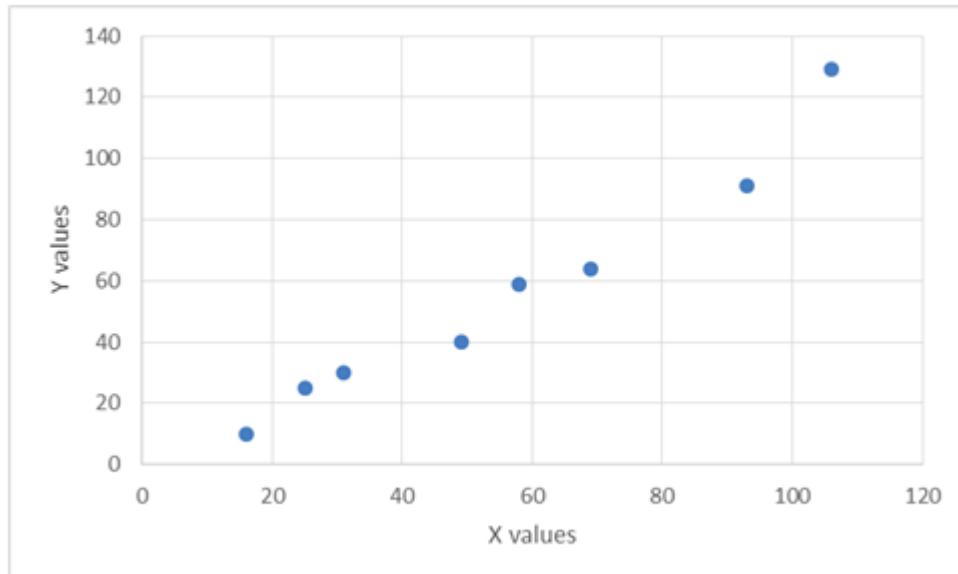
0.991

Correct Answers

Between 0.99 and 0.992

**Question 10****1 / 1 pts**

	X	Y
A	16	10
B	25	25
C	31	30
D	49	40
E	58	59
F	69	64
G	93	91
H	106	129



Select the most similar correlation between X and Y

0.5

-0.5 -1 1 0**Correct!**

# Knowledge Check: Review of Initial Data Exploration Techniques

<b>Due</b> No due date	<b>Points</b> 2	<b>Questions</b> 2
<b>Time Limit</b> None	<b>Allowed Attempts</b> Unlimited	

## Instructions

2 questions.

LO: Review and summarize data exploration techniques for use in initial data analysis

[Take the Quiz Again](#)

## Attempt History

	<b>Attempt</b>	<b>Time</b>	<b>Score</b>
<b>KEPT</b>	<a href="#"><u>Attempt 2</u></a>	less than 1 minute	2 out of 2
<b>LATEST</b>	<a href="#"><u>Attempt 2</u></a>	less than 1 minute	2 out of 2
	<a href="#"><u>Attempt 1</u></a>	1 minute	0.67 out of 2

Submitted Oct 20 at 12:54am

### Question 1

1 / 1 pts

Select all correct sentences that make up the similarity property.

Correct!



Cosine and Jaccard-Coefficient are typical similarity measurement methods

**Correct!** Similarity is always in the range 0 to 1 If  $p = q$ , similarity,  $s(p, q)$  is one (or maximum similarity)**Correct!** Similarity is usually in the range -1 to 1

Correct. A and B are similarity property. Euclidean, Cosine, and Jaccard are common methods for similarity measurement.

## Question 2

1 / 1 pts

Select all correct sentences that define correlation.

Correlation 1 means that the vectors have a perfect negative linear relationship.

Correlation 0 means that the vectors have a perfect positive linear relationship.

**Correct!**

Correlation -1 means that the vectors have a perfect negative linear relationship.

Correct. If the correlation is -1, the vectors follow the perfect negative linear correlation.

# Knowledge Check: Data Attributes Needed for Data Mining

**Due** No due date    **Points** 2    **Questions** 2  
**Time Limit** None    **Allowed Attempts** Unlimited

## Instructions

2 questions!

LO: Recognize attributes of data needed for data mining

[Take the Quiz Again](#)

## Attempt History

	Attempt	Time	Score
KEPT	<a href="#">Attempt 4</a>	less than 1 minute	2 out of 2
LATEST	<a href="#">Attempt 4</a>	less than 1 minute	2 out of 2
	<a href="#">Attempt 3</a>	less than 1 minute	1 out of 2
	<a href="#">Attempt 2</a>	less than 1 minute	2 out of 2
	<a href="#">Attempt 1</a>	less than 1 minute	1 out of 2

Submitted Oct 24 at 7:07pm

### Question 1

1 / 1 pts

Select all correct sentences that define attributes.

Correct!



The discrete attribute has only a finite or countably infinite set of values.

A discrete attribute can be mapped to a countably infinite set of values!

- The binary attributes are discrete attributes.
- The continuous attributes are often represented as integer variables.

## Question 2

1 / 1 pts

Select from the list below the attributes that could contribute to data quality problems.

Correct!

- Missing values

Missing values and duplicate data are intuitively critical to data quality.

Correct!

- Outliers

Outliers are data objects with characteristics that are considerably different than most of the other data objects in the data set.

Correct!

- Duplicate Data

Missing values and duplicate data are intuitively critical to data quality.

Correct!

- Noise

Noise refers to the modification of original values.

Noise refers to the modification of original values. Outliers are data objects with characteristics that are considerably different than most of the other data objects in the data set. Missing values and duplicate data are intuitively critical to data quality.

# Knowledge Check: Introduction to Data Mining Tasks

**Due** No due date    **Points** 2    **Questions** 2  
**Time Limit** None    **Allowed Attempts** Unlimited

## Instructions

2 questions!

LO: Describe different data mining tasks

[Take the Quiz Again](#)

## Attempt History

	<b>Attempt</b>	<b>Time</b>	<b>Score</b>
KEPT	<a href="#">Attempt 2</a>	less than 1 minute	2 out of 2
LATEST	<a href="#">Attempt 2</a>	less than 1 minute	2 out of 2
	<a href="#">Attempt 1</a>	less than 1 minute	0.5 out of 2

Submitted Oct 20 at 12:50am

### Question 1

1 / 1 pts

Select all supervised learning methodologies

Clustering

Dimensionality Reduction

Correct!

Classification or categorization

Classification is a discrete data based supervised learning method.

**Correct!**

- 
- Regression

Regression is a continuous data based supervised learning method.

**Correct!****Question 2****1 / 1 pts**

Select all unsupervised learning applications

**Correct!**

- 
- Marketing and sales promotion using Association Rule

Association Rule Mining is an unsupervised learning method.

**Correct!**

- 
- Document clustering

Clustering is an unsupervised learning method.

- 
- Face recognition

- 
- Fraud detection in credit card transactions

# Knowledge Check: History and Purpose of Data Mining

**Due** No due date    **Points** 2    **Questions** 2  
**Time Limit** None    **Allowed Attempts** Unlimited

## Instructions

2 questions!

LO: Explain the history and purpose of data mining across multiple disciplines

[Take the Quiz Again](#)

## Attempt History

	Attempt	Time	Score
KEPT	<a href="#">Attempt 2</a>	less than 1 minute	2 out of 2
LATEST	<a href="#">Attempt 2</a>	less than 1 minute	2 out of 2
	<a href="#">Attempt 1</a>	less than 1 minute	1 out of 2

Submitted Oct 19 at 10:03pm

### Question 1

1 / 1 pts

Discovering a student name in the class roster is an example of data mining function.

True

False

Correct!

Data mining is to discover (or extract) meaningful, previously unknown, and potentially useful information (or patterns) through exploring and analyzing large quantities of data.

**Question 2****1 / 1 pts**

Big data is required for data mining.

 True False**Correct!**

Data mining can be performed with small data.

# Knowledge Check: History and Purpose of Data Mining

**Due** No due date    **Points** 2    **Questions** 2  
**Time Limit** None    **Allowed Attempts** Unlimited

## Instructions

2 questions!

LO: Explain the history and purpose of data mining across multiple disciplines

[Take the Quiz Again](#)

## Attempt History

	Attempt	Time	Score
KEPT	<a href="#">Attempt 2</a>	less than 1 minute	2 out of 2
LATEST	<a href="#">Attempt 2</a>	less than 1 minute	2 out of 2
	<a href="#">Attempt 1</a>	less than 1 minute	1 out of 2

Submitted Oct 19 at 10:03pm

### Question 1

1 / 1 pts

Discovering a student name in the class roster is an example of data mining function.

True

False

Correct!

Data mining is to discover (or extract) meaningful, previously unknown, and potentially useful information (or patterns) through exploring and analyzing large quantities of data.

**Question 2**

1 / 1 pts

Big data is required for data mining.

 True False**Correct!**

Data mining can be performed with small data.

# Final Exam

**Due** Dec 8 at 1:15am      **Points** 40      **Questions** 20

**Available** Dec 6 at 12am - Dec 8 at 2am 2 days

**Time Limit** 75 Minutes

## Instructions

No collaboration

75 mins

## Attempt History

	<b>Attempt</b>	<b>Time</b>	<b>Score</b>
<b>LATEST</b>	<a href="#"><u>Attempt 1</u></a>	50 minutes	38 out of 40

! Correct answers are hidden.

Score for this quiz: **38** out of 40

Submitted Dec 6 at 2:37pm

This attempt took 50 minutes.

### Question 1

2 / 2 pts

Example of market basket transactions.

Transaction ID Items Bought

1 {a, b, d, e}

2 {b, c, d}

3 {a, b, d, e}

- 4      {a, c, d, e}
- 5      {b, c, d, e}
- 6      {b, d, e}
- 7      {c, d}
- 8      {a, b, c}
- 9      {a, d, e}
- 10     {b, d}

Compute the support for itemset {b,c,d}.

0.2

## Question 2

2 / 2 pts

Example of market basket transactions.

### Transaction ID Items Bought

- | Transaction ID | Items Bought |
|----------------|--------------|
| 1              | {a, b, d, e} |
| 2              | {b, c, d}    |
| 3              | {a, b, d, e} |
| 4              | {a, c, d, e} |

5      {b, c, d, e}

6      {b, d, e}

7      {c, d}

8      {a, b, c}

9      {a, d, e}

10     {b, d}

Compute the confidence for the association rules  $\{a, e\} \rightarrow \{b\}$ .

0.5

### Question 3

2 / 2 pts

Transaction ID Items Bought

1      {Milk, Beer, Diapers}

2      {Bread, Butter, Milk}

3      {Milk, Diapers, Cookies}

4      {Bread, Butter, Cookies}

5      {Beer, Cookies, Diapers}

6      {Milk, Diapers, Bread, Butter}

- 7 {Bread, Butter, Diapers}
- 8 {Beer, Diapers}
- 9 {Milk, Diapers, Bread, Butter}
- 10 {Beer, Cookies}

What is the confidence of the rule Beer  $\rightarrow$  Cookies

0.5

#### Question 4

2 / 2 pts

Example of market basket transactions.

#### Transaction ID Items Bought

- 1 {a, b, d, e}
- 2 {b, c, d}
- 3 {a, b, d, e}
- 4 {a, c, d, e}
- 5 {b, c, d, e}
- 6 {b, d, e}
- 7 {c, d}

8            {a, b, c}

9            {a, d, e}

10          {b, d}

If the minimum support is set at 40%, how many frequent 3-itemsets will be found? Note that if a 3 itemset is a subset of a larger itemset, it counts as one occurrence.

2

### Question 5

2 / 2 pts

Transaction ID Items Bought

1            {Milk, Beer, Diapers}

2            {Bread, Butter, Milk}

3            {Milk, Diapers, Cookies}

4            {Bread, Butter, Cookies}

5            {Beer, Cookies, Diapers}

6            {Milk, Diapers, Bread, Butter}

7            {Bread, Butter, Diapers}

8            {Beer, Diapers}

9 {Milk, Diapers, Bread, Butter}

10 {Beer, Cookies}

What is the maximum number of size 3-itemsets that can be derived from this data set. Note that the itemsets derived may not be in the dataset.

20

35

40

15

None of the above.

### Question 6

2 / 2 pts

Which of the following is an application of autoencoders?

feature learning.

dimensionality reduction.

Rule mining

all of the above.

### Question 7

2 / 2 pts

The number of neurons in the output layer should match the number of classes (Where the number of classes is greater than 2) in a supervised learning task.

---

True

---

False

**Question 8****2 / 2 pts**

What is an autoencoder?

---

a neural network that copies its input to its output.

---

a neural network that codes itself.

---

a neural network that maps an output to an input through a hidden layer.

---

a neural network that is trained to attempt to copy its input to its output.

**Question 9****2 / 2 pts**

Consider a movie recommendation application, where an user can log in and select movies they like based on their previous preferences.

Which neural network architecture would be suitable to complete this task

---

Fully-Connected Neural Network.

---

Convolutional Neural Network.

---

- Recurrent Neural Network.
- 
- Restricted Boltzmann Machine.

Incorrect

**Question 10****0 / 2 pts**

Increase in size of a convolutional hidden layer would not necessarily increase the performance of a convolutional network.

- 
- True
- 
- False

**Question 11****2 / 2 pts**

Select different aspects of cluster validation.

- 
- Determining the clustering tendency of a set of data
- 
- To improve the complexity of the algorithm
- 
- Determining the 'correct' number of clusters
- 
- None of the above

correct these are some of the aspects of cluster validation!

**Question 12****2 / 2 pts**

What is the purpose of cluster analysis?

- To avoid finding patterns in noise
- To compare clustering algorithms
- To compare two sets of clusters
- all of the above
- None of the above

we do cluster validity to avoid clustering noise and find a suitable algorithm for our data

**Question 13****2 / 2 pts**

Select qualities of clusters produced by a good clustering algorithm.

- Intra-cluster distances are minimized
- Inter-cluster distances are maximized
- Number of clusters produced
- A and B are correct

good cluster have strong cohesion within the cluster and maximum distance between them.

**Question 14****2 / 2 pts**

Select all statements that are true for Hierarchical clustering.

Although it requires high storage, high computing is not required

Robust to noise

Robust to outliers

It is not easy to decide on the number of cluster because the complicated dendrogram disturbs this decision

None of above

**Question 15****2 / 2 pts**

Review the following table.

Data	X	Y
A	1	2
B	2	2
C	2	3
D	8	7
E	8	8
F	8	2

Select the correct number of core points and DBSCAN clustering result with the parameters below.

Eps: 5

Minimum samples (MinPts):3

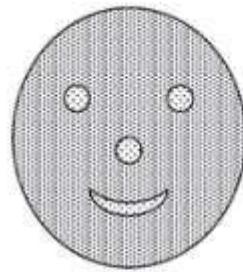
## Distance measurement: Euclidean distance

- 5, cluster1: {A,B,C}, cluster2: {D,E,F}
- 6, cluster1: {A,B}, cluster2: {C,D,E}, cluster3: {F}
- 4, cluster1: {A,B,C}, cluster2: {D, E, F}
- 5, cluster1: {A}, cluster2: {B}, cluster3: {C}, cluster4: {D}, cluster5: {E}
- None of above

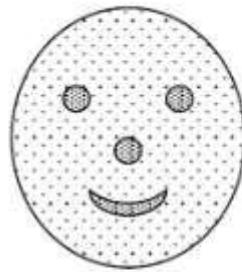
**Question 16**

2 / 2 pts

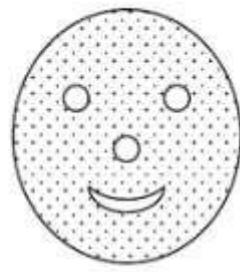
Review the image below.



(a)



(b)



(c)



(d)

**Darkness or number of dots represent density. Lines are used only to distinguish regions and do not represent points.**

If you want to find the patterns represented by the nose, eyes, and mouth using k-Means clustering, select all figures that be well-clustered?

 (a) (b) (c)

(d) None of above**Question 17****2 / 2 pts**

Select all that are not true about DBSCAN.

- it has trouble when the clusters have widely varying densities
- it is very sensitive to noise and to the size of the data
- DBSCAN has a complexity of order  $O(n^3)$
- it can be expensive when the computation of nearest neighbors requires computing all pairwise proximities
- A and B are correct
- B and C are correct

DBSCAN is relatively resistant to noise and can handle clusters of arbitrary shapes and sizes. It has complexity order  $O(n \log n)$

- A, C and D are correct

**Question 18****2 / 2 pts**

Review the table below.

Instance	X-Axis	Class
i	-3	-
ii	-2	-
iii	-1	+
iv	0	+
v	1	+
vi	2	-

Select a correct hyper-plane equation of the data below using Non-linear SVM.

## Feature Map ( $\{R^1 \rightarrow R^2\}$ )

Y=1

Y=1.5

Y=2

Y=2.5

None of above

### Question 19

2 / 2 pts

Review the table below.

Instance	X-Axis	Class
i	-3	-
ii	-2	-
iii	-1	+
iv	0	+
v	1	+
vi	2	-

Select all the support vector instances using Non-linear SVM

## Feature Map ( $\{R^1 \rightarrow R^2\}$ )

i

ii

iii

iv

v

vi

### Question 20

2 / 2 pts

Review the table below.

A	B	C	Class
F	F	F	Y
F	F	T	N
F	T	T	N
F	T	T	N
F	F	T	Y
T	F	T	Y
T	F	T	N
T	F	T	N
T	T	T	Y
T	F	T	Y

Select the correct prediction class for a test sample (A=F, B=T, C=F) using the naïve Bayes approach.

'Y'

'N'

Anyone

Quiz Score: **38** out of 40

# Final Exam

**Due** Dec 8 at 1:15am      **Points** 40      **Questions** 20

**Available** Dec 6 at 12am - Dec 8 at 2am 2 days

**Time Limit** 75 Minutes

## Instructions

No collaboration

75 mins

## Attempt History

	<b>Attempt</b>	<b>Time</b>	<b>Score</b>
<b>LATEST</b>	<a href="#"><u>Attempt 1</u></a>	50 minutes	38 out of 40

! Correct answers are hidden.

Score for this quiz: **38** out of 40

Submitted Dec 6 at 2:37pm

This attempt took 50 minutes.

<b>Question 1</b>		<b>2 / 2 pts</b>
Example of market basket transactions.		
Transaction ID Items Bought		
1	{a, b, d, e}	
2	{b, c, d}	
3	{a, b, d, e}	

- 4      {a, c, d, e}
- 5      {b, c, d, e}
- 6      {b, d, e}
- 7      {c, d}
- 8      {a, b, c}
- 9      {a, d, e}
- 10     {b, d}

Compute the support for itemset {b,c,d}.

0.2

## Question 2

2 / 2 pts

Example of market basket transactions.

### Transaction ID Items Bought

- 1      {a, b, d, e}
- 2      {b, c, d}
- 3      {a, b, d, e}
- 4      {a, c, d, e}

5      {b, c, d, e}

6      {b, d, e}

7      {c, d}

8      {a, b, c}

9      {a, d, e}

10     {b, d}

Compute the confidence for the association rules  $\{a, e\} \rightarrow \{b\}$ .

0.5

### Question 3

2 / 2 pts

Transaction ID Items Bought

1      {Milk, Beer, Diapers}

2      {Bread, Butter, Milk}

3      {Milk, Diapers, Cookies}

4      {Bread, Butter, Cookies}

5      {Beer, Cookies, Diapers}

6      {Milk, Diapers, Bread, Butter}

- 7      {Bread, Butter, Diapers}
- 8      {Beer, Diapers}
- 9      {Milk, Diapers, Bread, Butter}
- 10     {Beer, Cookies}

What is the confidence of the rule Beer  $\rightarrow$  Cookies

0.5

#### Question 4

2 / 2 pts

Example of market basket transactions.

#### Transaction ID Items Bought

- 1      {a, b, d, e}
- 2      {b, c, d}
- 3      {a, b, d, e}
- 4      {a, c, d, e}
- 5      {b, c, d, e}
- 6      {b, d, e}
- 7      {c, d}

8            {a, b, c}

9            {a, d, e}

10          {b, d}

If the minimum support is set at 40%, how many frequent 3-itemsets will be found? Note that if a 3 itemset is a subset of a larger itemset, it counts as one occurrence.

2

### Question 5

2 / 2 pts

Transaction ID Items Bought

1            {Milk, Beer, Diapers}

2            {Bread, Butter, Milk}

3            {Milk, Diapers, Cookies}

4            {Bread, Butter, Cookies}

5            {Beer, Cookies, Diapers}

6            {Milk, Diapers, Bread, Butter}

7            {Bread, Butter, Diapers}

8            {Beer, Diapers}

9 {Milk, Diapers, Bread, Butter}

10 {Beer, Cookies}

What is the maximum number of size 3-itemsets that can be derived from this data set. Note that the itemsets derived may not be in the dataset.

20

35

40

15

None of the above.

### Question 6

2 / 2 pts

Which of the following is an application of autoencoders?

feature learning.

dimensionality reduction.

Rule mining

all of the above.

### Question 7

2 / 2 pts

The number of neurons in the output layer should match the number of classes (Where the number of classes is greater than 2) in a supervised learning task.

---

True

---

False

**Question 8****2 / 2 pts**

What is an autoencoder?

---

a neural network that copies its input to its output.

---

a neural network that codes itself.

---

a neural network that maps an output to an input through a hidden layer.

---

a neural network that is trained to attempt to copy its input to its output.

**Question 9****2 / 2 pts**

Consider a movie recommendation application, where an user can log in and select movies they like based on their previous preferences.

Which neural network architecture would be suitable to complete this task

---

Fully-Connected Neural Network.

---

Convolutional Neural Network.

---

- Recurrent Neural Network.
- 
- Restricted Boltzmann Machine.

Incorrect

**Question 10****0 / 2 pts**

Increase in size of a convolutional hidden layer would not necessarily increase the performance of a convolutional network.

- 
- True
- 
- False

**Question 11****2 / 2 pts**

Select different aspects of cluster validation.

- 
- Determining the clustering tendency of a set of data
- 
- To improve the complexity of the algorithm
- 
- Determining the 'correct' number of clusters
- 
- None of the above

correct these are some of the aspects of cluster validation!

**Question 12****2 / 2 pts**

What is the purpose of cluster analysis?

- To avoid finding patterns in noise
- To compare clustering algorithms
- To compare two sets of clusters
- all of the above
- None of the above

we do cluster validity to avoid clustering noise and find a suitable algorithm for our data

**Question 13****2 / 2 pts**

Select qualities of clusters produced by a good clustering algorithm.

- Intra-cluster distances are minimized
- Inter-cluster distances are maximized
- Number of clusters produced
- A and B are correct

good cluster have strong cohesion within the cluster and maximum distance between them.

**Question 14****2 / 2 pts**

Select all statements that are true for Hierarchical clustering.

Although it requires high storage, high computing is not required

Robust to noise

Robust to outliers

It is not easy to decide on the number of cluster because the complicated dendrogram disturbs this decision

None of above

**Question 15****2 / 2 pts**

Review the following table.

Data	X	Y
A	1	2
B	2	2
C	2	3
D	8	7
E	8	8
F	8	2

Select the correct number of core points and DBSCAN clustering result with the parameters below.

Eps: 5

Minimum samples (MinPts):3

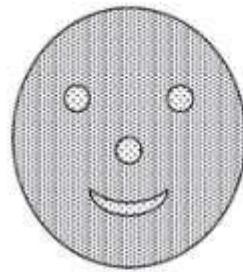
### Distance measurement: Euclidean distance

- 5, cluster1: {A,B,C}, cluster2: {D,E,F}
- 6, cluster1: {A,B}, cluster2: {C,D,E}, cluster3: {F}
- 4, cluster1: {A,B,C}, cluster2: {D, E, F}
- 5, cluster1: {A}, cluster2: {B}, cluster3: {C}, cluster4: {D}, cluster5: {E}
- None of above

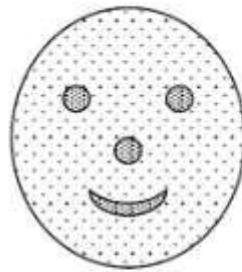
### Question 16

2 / 2 pts

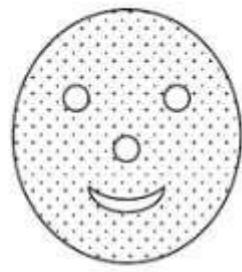
Review the image below.



(a)



(b)



(c)



(d)

**Darkness or number of dots represent density. Lines are used only to distinguish regions and do not represent points.**

If you want to find the patterns represented by the nose, eyes, and mouth using k-Means clustering, select all figures that be well-clustered?

 (a) (b) (c)

(d) None of above**Question 17****2 / 2 pts**

Select all that are not true about DBSCAN.

- it has trouble when the clusters have widely varying densities
- it is very sensitive to noise and to the size of the data
- DBSCAN has a complexity of order  $O(n^3)$
- 
- it can be expensive when the computation of nearest neighbors requires computing all pairwise proximities
- A and B are correct
- B and C are correct

DBSCAN is relatively resistant to noise and can handle clusters of arbitrary shapes and sizes. It has complexity order  $O(n \log n)$

- A, C and D are correct

**Question 18****2 / 2 pts**

Review the table below.

Instance	X-Axis	Class
i	-3	-
ii	-2	-
iii	-1	+
iv	0	+
v	1	+
vi	2	-

Select a correct hyper-plane equation of the data below using Non-linear SVM.

## Feature Map ( $\{R^1 \rightarrow R^2\}$ )

Y=1

Y=1.5

Y=2

Y=2.5

None of above

### Question 19

2 / 2 pts

Review the table below.

Instance	X-Axis	Class
i	-3	-
ii	-2	-
iii	-1	+
iv	0	+
v	1	+
vi	2	-

Select all the support vector instances using Non-linear SVM

## Feature Map ( $\{R^1 \rightarrow R^2\}$ )

i

ii

iii

iv

v

vi

### Question 20

2 / 2 pts

Review the table below.

A	B	C	Class
F	F	F	Y
F	F	T	N
F	T	T	N
F	T	T	N
F	F	T	Y
T	F	T	Y
T	F	T	N
T	F	T	N
T	T	T	Y
T	F	T	Y

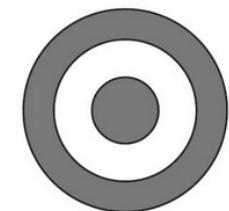
Select the correct prediction class for a test sample (A=F, B=T, C=F) using the naïve Bayes approach.

'Y'

'N'

Anyone

Quiz Score: **38** out of 40

<b>Question 1</b>	2 / 2 pts	<b>Question 2</b>	2 / 2 pts	<b>Question 3</b>	2 / 2 pts
<p>Example of market basket transactions.</p> <p>Transaction ID Items Bought</p> <p>1 {a, b, d, e} 2 {b, c, d} 3 {a, b, d, e} 4 {a, c, d, e} 5 {b, c, d, e} 6 {b, d, e} 7 {c, d} 8 {a, b, c} 9 {a, d, e} 10 {b, d}</p> <p>Compute the confidence for the association rules {a, e} -&gt; {b}.</p> <div style="border: 1px solid #ccc; padding: 2px; width: fit-content;">0.5</div>					
<p>Transaction ID Items Bought</p> <p>1 {Milk, Beer, Diapers} 2 {Bread, Butter, Milk} 3 {Milk, Diapers, Cookies} 4 {Bread, Butter, Cookies} 5 {Beer, Cookies, Diapers} 6 {Milk, Diapers, Bread, Butter} 7 {Bread, Butter, Diapers} 8 {Beer, Diapers} 9 {Milk, Diapers, Bread, Butter} 10 {Beer, Cookies}</p> <p>What is the confidence of the rule Beer -&gt; Cookies</p> <div style="border: 1px solid #ccc; padding: 2px; width: fit-content;">0.5</div>					
<p>Transaction ID Items Bought</p> <p>1 {Milk, Beer, Diapers} 2 {Bread, Butter, Milk} 3 {Milk, Diapers, Cookies} 4 {Bread, Butter, Cookies} 5 {Beer, Cookies, Diapers} 6 {Milk, Diapers, Bread, Butter} 7 {Bread, Butter, Diapers} 8 {Beer, Diapers} 9 {Milk, Diapers, Bread, Butter} 10 {Beer, Cookies}</p> <p>Compute the support for itemset {b} -&gt; {c,d}.</p> <div style="border: 1px solid #ccc; padding: 2px; width: fit-content;">0.2</div>					
<b>Question 4</b>	2 / 2 pts	<b>Question 5</b>	2 / 2 pts	<b>Question 8</b>	2 / 2 pts
<p>Transaction ID Items Bought</p> <p>1 {Milk, Beer, Diapers} 2 {Bread, Butter, Milk} 3 {Milk, Diapers, Cookies} 4 {Bread, Butter, Cookies} 5 {Beer, Cookies, Diapers} 6 {Milk, Diapers, Bread, Butter} 7 {Bread, Butter, Diapers} 8 {Beer, Diapers} 9 {Milk, Diapers, Bread, Butter} 10 {Beer, Cookies}</p> <p>Apriori pruning based on support is a greedy strategy</p> <div style="border: 1px solid #ccc; padding: 2px; width: fit-content;"> <input type="radio"/> True  <input checked="" type="radio"/> False     </div>					
<p><b>Question 6</b></p>					
<p>Select the true statements about neural network (NN) complexity.</p> <div style="border: 1px solid #ccc; padding: 2px; width: fit-content;"> <input type="checkbox"/> CNN is always more complex than DNN  <input checked="" type="checkbox"/> Adding more hidden layers increases complexity of a NN  <input type="checkbox"/> An Autoencoder is always more complex than a CNN  <input checked="" type="checkbox"/> Complexity metric defined for a NN may not make sense for other classification engines     </div>					
<p><b>Question 7</b></p>					
<p>What is an autoencoder?</p> <div style="border: 1px solid #ccc; padding: 2px; width: fit-content;"> <input type="radio"/> a neural network that copies its input to its output.  <input type="radio"/> a neural network that codes itself.  <input type="radio"/> a neural network that maps an output to an input through a hidden layer.  <input checked="" type="radio"/> a neural network that is trained to attempt to copy its input to its output.     </div>					
<p><b>Question 9</b></p>					
<p>Consider a movie recommendation application, where an user can log in and select movies they like based on their previous preferences.</p> <p>Which neural network architecture would be suitable to complete this task</p> <div style="border: 1px solid #ccc; padding: 2px; width: fit-content;"> <input type="radio"/> Fully-Connected Neural Network.  <input type="radio"/> Convolutional Neural Network.  <input type="radio"/> Recurrent Neural Network.  <input checked="" type="radio"/> Restricted Boltzmann Machine.     </div>					
<p><b>Question 10</b></p>					
<p>Increase in size of a convolutional hidden layer would not necessarily increase the performance of a convolutional network.</p> <div style="border: 1px solid #ccc; padding: 2px; width: fit-content;"> <input type="radio"/> True  <input checked="" type="radio"/> False     </div>					
<b>Question 13</b>	2 / 2 pts	<b>Question 14</b>	2 / 2 pts	<b>Question 15</b>	2 / 2 pts
<p>What is the purpose of cluster analysis?</p> <div style="border: 1px solid #ccc; padding: 2px; width: fit-content;"> <input type="radio"/> To avoid finding patterns in noise  <input type="radio"/> To compare clustering algorithms  <input type="radio"/> To compare two sets of clusters  <input checked="" type="radio"/> all of the above  <input type="radio"/> None of the above     </div> <div style="border: 1px solid #ccc; padding: 2px; width: fit-content; margin-top: 10px;">we do cluster validity to avoid clustering noise and find a suitable algorithm for our data</div>					
<p>Select different aspects of cluster validation.</p> <div style="border: 1px solid #ccc; padding: 2px; width: fit-content;"> <input checked="" type="checkbox"/> Determining the clustering tendency of a set of data  <input type="checkbox"/> To improve the complexity of the algorithm  <input checked="" type="checkbox"/> Determining the 'correct' number of clusters  <input type="checkbox"/> None of the above     </div> <div style="border: 1px solid #ccc; padding: 2px; width: fit-content; margin-top: 10px;">correct these are some of the aspects of cluster validation!</div>					
<p>Identify the number of clusters using the center-based, contiguity-based, and density-based definitions for below figure.</p>  <div style="border: 1px solid #ccc; padding: 2px; width: fit-content;"> <input checked="" type="radio"/> center-based: 1 cluster, contiguity-based: 2 clusters, density-based: 2 cluster  <input type="radio"/> center-based: 2 cluster, contiguity-based: 2 clusters, density-based: 1 cluster  <input type="radio"/> center-based: 3 cluster, contiguity-based: 1 clusters, density-based: 3 cluster  <input type="radio"/> center-based: 2 cluster, contiguity-based: 2 clusters, density-based: 2 cluster  <input type="radio"/> None of above     </div>					

# Final Exam

**Due** Dec 9 at 1:15am      **Points** 40      **Questions** 20

**Available** Dec 6 at 12am - Dec 9 at 2am 3 days

**Time Limit** 75 Minutes

## Instructions

No collaboration

75 mins

## Attempt History

	<b>Attempt</b>	<b>Time</b>	<b>Score</b>
<b>LATEST</b>	<a href="#"><u>Attempt 1</u></a>	74 minutes	20 out of 40

! Correct answers are hidden.

Score for this quiz: **20** out of 40

Submitted Dec 8 at 1:51pm

This attempt took 74 minutes.

Incorrect	<b>Question 1</b>	<b>0 / 2 pts</b>
Which of the above statements are true for any A, B, and C?		
<input type="radio"/> If A $\rightarrow$ B then B $\rightarrow$ A.		
<input type="radio"/> If A $\rightarrow$ B and B $\rightarrow$ C then A $\rightarrow$ C.		
<input type="radio"/> If A $\rightarrow$ C then A union B $\rightarrow$ C.		
<input type="radio"/> If A union B $\rightarrow$ C then A $\rightarrow$ C.		
<input type="radio"/> I & II		

I, II, & III I, II, & IV II & III none**Question 2****2 / 2 pts**

Example of market basket transactions.

Transaction ID Items Bought

1 {a, b, d, e}

2 {b, c, d}

3 {a, b, d, e}

4 {a, c, d, e}

5 {b, c, d, e}

6 {b, d, e}

7 {c, d}

8 {a, b, c}

9 {a, d, e}

10 {b, d}

Compute the confidence for the association rules  $\{a, e\} \rightarrow \{b\}$ .

0.5

**Question 3****2 / 2 pts**

Example of market basket transactions.

Transaction ID Items Bought

1             $\{a, b, d, e\}$

2             $\{b, c, d\}$

3             $\{a, b, d, e\}$

4             $\{a, c, d, e\}$

5             $\{b, c, d, e\}$

6             $\{b, d, e\}$

7             $\{c, d\}$

8             $\{a, b, c\}$

9             $\{a, d, e\}$

10           $\{b, d\}$

Compute the support for itemset  $\{b, c, d\}$ .

0.2

Incorrect

**Question 4**

0 / 2 pts

Example of market basket transactions.

Transaction ID Items Bought

- |    |              |
|----|--------------|
| 1  | {a, b, d, e} |
| 2  | {b, c, d}    |
| 3  | {a, b, d, e} |
| 4  | {a, c, d, e} |
| 5  | {b, c, d, e} |
| 6  | {b, d, e}    |
| 7  | {c, d}       |
| 8  | {a, b, c}    |
| 9  | {a, d, e}    |
| 10 | {b, d}       |

If the minimum support is set at 40%, how many frequent 3-itemsets will be found? Note that if a 3 itemset is a subset of a larger itemset, it counts as one occurrence.

4

**Question 5****2 / 2 pts**

Transaction ID Items Bought

- 1 {Milk, Beer, Diapers}
- 2 {Bread, Butter, Milk}
- 3 {Milk, Diapers, Cookies}
- 4 {Bread, Butter, Cookies}
- 5 {Beer, Cookies, Diapers}
- 6 {Milk, Diapers, Bread, Butter}
- 7 {Bread, Butter, Diapers}
- 8 {Beer, Diapers}
- 9 {Milk, Diapers, Bread, Butter}
- 10 {Beer, Cookies}

What is the maximum size of frequent itemsets that can be extracted  
(assuming minsup > 0.1)?

 10

9 4 3**Question 6****2 / 2 pts**

Increase in size of a convolutional hidden layer would not necessarily increase the performance of a convolutional network.

 True False**Question 7****2 / 2 pts**

Select the true statements about neural network (NN) complexity.

 CNN is always more complex than DNN Adding more hidden layers increases complexity of a NN An Autoencoder is always more complex than a CNN

Complexity metric defined for a NN may not make sense for other classification engines

Incorrect

**Question 8****0 / 2 pts**

Consider a movie recommendation application, where an user can log in and select movies they like based on their previous preferences.

Which neural network architecture would be suitable to complete this task

- 
- Fully-Connected Neural Network.
  - Convolutional Neural Network.
  - Recurrent Neural Network.
  - Restricted Boltzmann Machine.
- 

Incorrect

**Question 9****0 / 2 pts**

The number of neurons in the output layer should match the number of classes (Where the number of classes is greater than 2) in a supervised learning task.

- 
- True
  - False
- 

Partial

**Question 10****1 / 2 pts**

Select all that is true about Restricted Boltzman Machine (RBM)

- 
- Can be used for recommendation systems
-

- RBM training is often probabilistic
- RBM is a kind of Deep Belief Network
- None of the above.

Incorrect

**Question 11**

0 / 2 pts

Review the table below.

Cluster	Tennis	Golf	Baseball	Basketball	Football	Hockey	Total	Entropy	Purity
#1	1	1	0	11	4	676	693		
#2	27	89	333	827	253	33	1562		
#3	326	465	8	105	16	29	949		
Total	354	555	341	943	273	738	3204		

Select the collect entropy set of the confusion matrix.

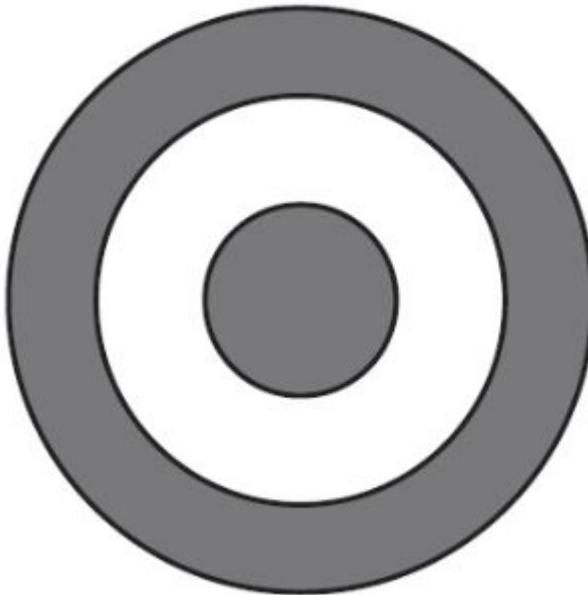
- Cluster #1: 1.84, Cluster #2: 0.2, Cluster #3: 1.7, Total: 1.44
- Cluster #1: 1.44, Cluster #2: 1.84, Cluster #3: 1.7, Total: 0.2
- Cluster #1: 0.2, Cluster #2: 1.7, Cluster #3: 1.84, Total: 1.44
- Cluster #1: 0.2, Cluster #2: 1.44, Cluster #3: 1.7, Total: 1.84
- None of above

Incorrect

**Question 12**

0 / 2 pts

Identify the number of clusters using the center-based, contiguity-based, and density-based definitions for below figure.



center-based: 1 cluster, contiguity-based: 2 clusters, density-based: 2 cluster



center-based: 2 cluster, contiguity-based: 2 clusters, density-based: 1 cluster



center-based: 3 cluster, contiguity-based: 1 clusters, density-based: 3 cluster



center-based: 2 cluster, contiguity-based: 2 clusters, density-based: 2 cluster



None of above

**Question 13****2 / 2 pts**

Select all different type of clusters.

- Well-separated
- Center-based
- Contiguous
- Density-based
- All of the above

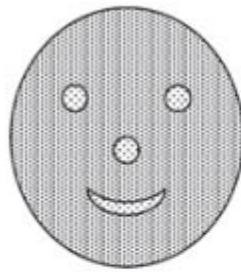
all are type of clustering

Partial

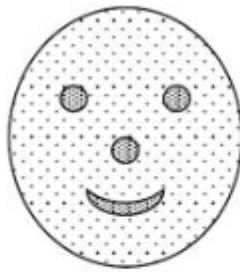
**Question 14**

1 / 2 pts

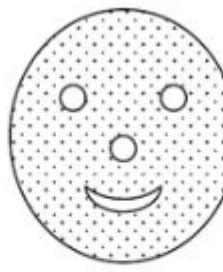
Review the image below.



(a)



(b)



(c)



(d)

**Darkness or number of dots represent density. Lines are used only to distinguish regions and do not represent points.**

If we want to find the patterns represented by the nose, eyes, and mouth using single linkage hierarchical clustering, select all figures that be well-clustered?

- (a)
- (b)

(c) (d) None of above

Incorrect

**Question 15**

0 / 2 pts

Review the following table.

Data	X	Y
x1	1	1
x2	6	6
x3	7	7
x4	3	2
x5	6	8
x6	2	1
x7	2	3

Select the correct pair of final centroid and clustering set using K-mean clustering with the parameters below.

K:2

Initial centroids: centroid1 = (4,3) and centroid2 = (9,7)

Distance measure: Euclidean distance



Cluster 1: (x1, x4, x6, x7) with centroid=(2, 1.75), Cluster 2: (x2, x3, x5) with centroid=(6.33, 7)



Cluster 1:  $(x_1, x_4, x_6, x_7)$  with centroid=(1.75, 2), Cluster 2:  $(x_2, x_3, x_5)$  with centroid=(6.33, 7)



Cluster 1:  $(x_1, x_4, x_6, x_7)$  with centroid=(2, 1.75), Cluster 2:  $(x_2, x_3, x_5)$  with centroid=(7, 6.33)



Cluster 1:  $(x_1, x_2, x_4, x_6, x_7)$  with centroid=(2, 1.75), Cluster 2:  $(x_3, x_5)$  with centroid=(6.33, 7)



Cluster 1:  $(x_1, x_2, x_4, x_6, x_7)$  with centroid=(2, 1.75), Cluster 2:  $(x_3, x_5)$  with centroid=(7, 6.33)

### Question 16

2 / 2 pts

Select all that are not true about DBSCAN.

it has trouble when the clusters have widely varying densities

it is very sensitive to noise and to the size of the data

DBSCAN has a complexity of order  $O(n^3)$



it can be expensive when the computation of nearest neighbors requires computing all pairwise proximities

A and B are correct

B and C are correct

DBSCAN is relatively resistant to noise and can handle clusters of arbitrary shapes and sizes. it has complexity order  $O(n \log n)$

- A, C and D are correct

### Question 17

2 / 2 pts

Select all statements that are true for Hierarchical clustering.

Although it requires high storage, high computing is not required

Robust to noise

Robust to outliers

It is not easy to decide on the number of cluster because the complicated dendrogram disturbs this decision

None of above

### Question 18

2 / 2 pts

Review the table below.

A	B	C	Class
F	F	F	Y
F	F	T	N
F	T	T	N
F	T	T	N
F	F	T	Y
T	F	T	Y
T	F	T	N
T	F	T	N
T	T	T	Y
T	F	T	Y

Select the correct prediction class for a test sample (A=F, B=T, C=F) using the naïve Bayes approach.

'Y'

'N'

Anyone

Incorrect

### Question 19

0 / 2 pts

Select the true statement for Nearest Neighbor classification

Although Nearest Neighbor classification has the data dimensionality issue, it can be solved by the scaling technique



k-NN classifier is a typical lazy learner because it spend so much time building the model even if the 'k' is very small



Drawing a circle is a mandatory task to find neighbor classes regardless of the number of 'K'



Determining the optimal value of 'K' is important for k-NN classifier performance.



None of above

Incorrect

**Question 20**

0 / 2 pts

Review the table below.

Instance	X-Axis	Class
i	-3	-
ii	-2	-
iii	-1	+
iv	0	+
v	1	+
vi	2	-

Select a correct hyper-plane equation of the data below using Non-linear SVM.

## Feature Map ( $\{R^1 \rightarrow R^2\}$ )

- Y=1
- Y=1.5
- Y=2
- Y=2.5
- None of above

Quiz Score: **20** out of 40

# Final Exam

**Due** Dec 9 at 1:15am      **Points** 40      **Questions** 20

**Available** Dec 6 at 12am - Dec 9 at 2am 3 days

**Time Limit** 75 Minutes

## Instructions

No collaboration

75 mins

## Attempt History

	Attempt	Time	Score
LATEST	<a href="#"><u>Attempt 1</u></a>	46 minutes	38 out of 40

! Correct answers are hidden.

Score for this quiz: **38** out of 40

Submitted Dec 8 at 4:43pm

This attempt took 46 minutes.

Question 1	2 / 2 pts
Example of market basket transactions.	
Transaction ID Items Bought	
1	{a, b, d, e}
2	{b, c, d}

- |    |              |
|----|--------------|
| 3  | {a, b, d, e} |
| 4  | {a, c, d, e} |
| 5  | {b, c, d, e} |
| 6  | {b, d, e}    |
| 7  | {c, d}       |
| 8  | {a, b, c}    |
| 9  | {a, d, e}    |
| 10 | {b, d}       |

If the minimum support is set at 40%, how many frequent 3-itemsets will be found? Note that if a 3 itemset is a subset of a larger itemset, it counts as one occurrence.

2

**Question 2****2 / 2 pts**

Select the reason why Apriori pruning in the search for frequent itemsets works.

- support count is monotonic with respect to itemsets.
- we search in transaction ID order.
- support count diverges as we add to the itemset.

- support count is anti-monotonic with respect to itemsets.
- it is an excellent heuristic, but it does not work 100% of the time.

**Question 3****2 / 2 pts**

Transaction ID Items Bought

- 1 {Milk, Beer, Diapers}
- 2 {Bread, Butter, Milk}
- 3 {Milk, Diapers, Cookies}
- 4 {Bread, Butter, Cookies}
- 5 {Beer, Cookies, Diapers}
- 6 {Milk, Diapers, Bread, Butter}
- 7 {Bread, Butter, Diapers}
- 8 {Beer, Diapers}
- 9 {Milk, Diapers, Bread, Butter}
- 10 {Beer, Cookies}

What is the maximum size of frequent itemsets that can be extracted  
(assuming minsup > 0.1)?

10 9 4 3**Question 4****2 / 2 pts**

Transaction ID Items Bought

- 1 {Milk, Beer, Diapers}
- 2 {Bread, Butter, Milk}
- 3 {Milk, Diapers, Cookies}
- 4 {Bread, Butter, Cookies}
- 5 {Beer, Cookies, Diapers}
- 6 {Milk, Diapers, Bread, Butter}
- 7 {Bread, Butter, Diapers}
- 8 {Beer, Diapers}
- 9 {Milk, Diapers, Bread, Butter}
- 10 {Beer, Cookies}

Find all itemsets (of size 3 or larger) that has the largest support.

{Milk, Beer, Diapers}

{Bread, Butter, Milk}

{Bread, Butter, Diapers}

{Milk, Diapers, Bread}

None of the above.

All of the above.

**Question 5****2 / 2 pts**

Which of the above statements are true for any A, B, and C?

- If  $A \rightarrow B$  then  $B \rightarrow A$ .
- If  $A \rightarrow B$  and  $B \rightarrow C$  then  $A \rightarrow C$ .
- If  $A \rightarrow C$  then  $A \cup B \rightarrow C$ .
- If  $A \cup B \rightarrow C$  then  $A \rightarrow C$ .
- I & II
- I, II, & III
- I, II, & IV
- II & III
- none

**Question 6****2 / 2 pts**

The number of neurons in the output layer should match the number of classes (Where the number of classes is greater than 2) in a supervised learning task.

- True
- False

**Question 7****2 / 2 pts**

Increase in size of a convolutional hidden layer would not necessarily increase the performance of a convolutional network.

True

False

### Question 8

2 / 2 pts

Is an autoencoder the same as performing Principal Component Analysis?

Not at all

No but conceptually they can be used for the same purpose

Autoencoders can do dimensionality reduction, but it is non-linear.

Yes they have the same purpose.

### Question 9

2 / 2 pts

Usage of deep learning is predicated with availability of large amounts of data. Is the data volume the only determining factor to check whether we can use deep learning ?

Yes, lower volumes lead to overfitting

No, balance among classes in the input data is the only determining factor volume does not matter



No, balance among classes in the input data is a co-factor along with the data volume



None of the above

### Question 10

2 / 2 pts

Which of the following is an application of autoencoders?

feature learning.

dimensionality reduction.

Rule mining

all of the above.

### Question 11

2 / 2 pts

Select all statements that are true for Hierarchical clustering.

Although it requires high storage, high computing is not required

Robust to noise

Robust to outliers

It is not easy to decide on the number of cluster because the complicated dendrogram disturbs this decision

- None of above

### Question 12

2 / 2 pts

Review the table below.

Cluster	Tennis	Golf	Baseball	Basketball	Football	Hockey	Total	Entropy	Purity
#1	1	1	0	11	4	676	693		
#2	27	89	333	827	253	33	1562		
#3	326	465	8	105	16	29	949		
Total	354	555	341	943	273	738	3204		

Select the collect purity set of the confusion matrix

- Cluster #1: 0.98, Cluster #2: 0.53, Cluster #3: 0.49, Total: 0.61

- Cluster #1: 0.53, Cluster #2: 0.98, Cluster #3: 0.61, Total: 0.49

- Cluster #1: 0.98, Cluster #2: 0.49, Cluster #3: 0.53, Total: 0.61

- Cluster #1: 0.53, Cluster #2: 0.49, Cluster #3: 0.61, Total: 0.98

- None of above

### Question 13

2 / 2 pts

Select all that are not true about DBSCAN.

- It has trouble when the clusters have widely varying densities

- it is very sensitive to noise and to the size of the data
- DBSCAN has a complexity of order  $O(n^3)$
- it can be expensive when the computation of nearest neighbors requires computing all pairwise proximities
- A and B are correct
- B and C are correct

DBSCAN is relatively resistant to noise and can handle clusters of arbitrary shapes and sizes. It has complexity order  $O(n \log n)$

- A, C and D are correct

### Question 14

2 / 2 pts

Review the following table.

Data	X	Y
x1	1	1
x2	6	6
x3	7	7
x4	3	2
x5	6	8
x6	2	1
x7	2	3

Select the correct pair of final centroid and clustering set using K-mean clustering with the parameters below.

K:2

Initial centroids: centroid1 = (4,3) and centroid2 = (9,7)

Distance measure: Euclidean distance



Cluster 1: (x1, x4, x6, x7) with centroid=(2, 1.75), Cluster 2: (x2, x3, x5) with centroid=(6.33, 7)



Cluster 1: (x1, x4, x6, x7) with centroid=(1.75, 2), Cluster 2: (x2, x3, x5) with centroid=(6.33, 7)



Cluster 1: (x1, x4, x6, x7) with centroid=(2, 1.75), Cluster 2: (x2, x3, x5) with centroid=(7, 6.33)



Cluster 1: (x1, x2, x4, x6, x7) with centroid=(2, 1.75), Cluster 2: (x3, x5) with centroid=(6.33, 7)



Cluster 1: (x1, x2, x4, x6, x7) with centroid=(2, 1.75), Cluster 2: (x3, x5) with centroid=(7, 6.33)

### Question 15

2 / 2 pts

Review the following table.

Data	X	Y
x1	1	1
x2	6	6
x3	7	7
x4	3	2
x5	6	8
x6	2	1
x7	2	3

Compute the sum of squared error (SSE) to evaluate K-mean clustering result with the parameters below.

K:2

Initial centroids: centroid1 = (4,3) and centroid2 = (9,7)

Distance measure: Euclidean distance

7.4167

Incorrect

### Question 16

0 / 2 pts

Review the following table.

Data	X	Y
x1	1	1
x2	6	6
x3	7	7
x4	3	2
x5	6	8
x6	2	1
x7	2	3

Compute the updated centroids after the first iteration using K-mean clustering with the parameters below.

K:2

Initial centroids: centroid1 = (4,3) and centroid2 = (9,7)

Distance measure: Euclidean

Type your answer in the same format as "centroid1 = (x1,y1) and centroid2 = (x2,y2)"

Each co-ordinate should be in the format A.BC, where A, B and C are integers. If you have 4 as one of the co-ordinates you should write 4.00. Please be mindful of the spaces and other formats.

centroid1 = (2.00,1.75) and centroid2 = (6.33,7.00)

### Question 17

2 / 2 pts

Review the following table.

	A	B	C	D	E
A	0	9	3	6	11
B	9	0	7	5	10
C	3	7	0	9	2
D	6	5	9	0	8
E	11	10	2	8	0

Select the correct DBSCAN clustering result with the parameters below

epsilon = 5, Minpts = 2

- cluster1: {A, C, E}, cluster2: {B,D}
- cluster1: {A}, cluster2: {B}, cluster3: {D}, and cluster4: {C, E}
- cluster1: {A, B}, cluster2: {C}, and cluster3: {D}, cluster4: {E}
- cluster1: {A}, cluster2: {B, E}, cluster3: {C}, and cluster4: {D}
- None of above

### Question 18

2 / 2 pts

Review the table below.

A	B	C	Class
F	F	F	Y
F	F	T	N
F	T	T	N
F	T	T	N
F	F	T	Y
T	F	T	Y
T	F	T	N
T	F	T	N
T	T	T	Y
T	F	T	Y

Select the correct prediction class for a test sample (A=F, B=T, C=F) using the naïve Bayes approach.

'Y'

'N'

Anyone

### Question 19

2 / 2 pts

Select the true statement for Nearest Neighbor classification

Although Nearest Neighbor classification has the data dimensionality issue, it can be solved by the scaling technique

k-NN classifier is a typical lazy learner because it spend so much time building the model even if the 'k' is very small

Drawing a circle is a mandatory task to find neighbor classes regardless of the number of 'K'

Determining the optimal value of 'K' is important for k-NN classifier performance.

None of above

### Question 20

2 / 2 pts

Review the table below.

Instance	X-Axis	Class
i	-3	-
ii	-2	-
iii	-1	+
iv	0	+
v	1	+
vi	2	-

Select all the support vector instances using Non-linear SVM

Feature Map ( $\{R^1 \rightarrow R^2\}$ )

 i

<input checked="" type="checkbox"/> ii
<input checked="" type="checkbox"/> iii
<input type="checkbox"/> iv
<input checked="" type="checkbox"/> v
<input checked="" type="checkbox"/> vi

Quiz Score: **38** out of 40

# Final Exam

**Due** Dec 9 at 1:15am      **Points** 40      **Questions** 20

**Available** Dec 6 at 12am - Dec 9 at 2am 3 days

**Time Limit** 75 Minutes

## Instructions

No collaboration

75 mins

## Attempt History

	<b>Attempt</b>	<b>Time</b>	<b>Score</b>
LATEST	<a href="#"><u>Attempt 1</u></a>	66 minutes	38 out of 40

! Correct answers are hidden.

Score for this quiz: **38** out of 40

Submitted Dec 8 at 3:52pm

This attempt took 66 minutes.

### Question 1

2 / 2 pts

Transaction ID Items Bought

- 1 {Milk, Beer, Diapers}
- 2 {Bread, Butter, Milk}
- 3 {Milk, Diapers, Cookies}
- 4 {Bread, Butter, Cookies}
- 5 {Beer, Cookies, Diapers}

- 6 {Milk, Diapers, Bread, Butter}
- 7 {Bread, Butter, Diapers}
- 8 {Beer, Diapers}
- 9 {Milk, Diapers, Bread, Butter}
- 10 {Beer, Cookies}

What is the maximum size of frequent itemsets that can be extracted  
(assuming minsup > 0.1)?

---

10

---

9

---

4

---

3

## Question 2

2 / 2 pts

Transaction ID Items Bought

- 1 {Milk, Beer, Diapers}
- 2 {Bread, Butter, Milk}
- 3 {Milk, Diapers, Cookies}
- 4 {Bread, Butter, Cookies}
- 5 {Beer, Cookies, Diapers}

- 6 {Milk, Diapers, Bread, Butter}
- 7 {Bread, Butter, Diapers}
- 8 {Beer, Diapers}
- 9 {Milk, Diapers, Bread, Butter}
- 10 {Beer, Cookies}

What is the confidence of the rule Beer -> Cookies

0.5

### Question 3

2 / 2 pts

Example of market basket transactions.

Transaction ID Items Bought

- 1 {a, b, d, e}
- 2 {b, c, d}
- 3 {a, b, d, e}
- 4 {a, c, d, e}
- 5 {b, c, d, e}
- 6 {b, d, e}
- 7 {c, d}
- 8 {a, b, c}

9            {a, d, e}

10          {b, d}

Compute the support for itemset {b,c,d}.

0.2

#### Question 4

2 / 2 pts

Example of market basket transactions.

Transaction ID Items Bought

1            {a, b, d, e}

2            {b, c, d}

3            {a, b, d, e}

4            {a, c, d, e}

5            {b, c, d, e}

6            {b, d, e}

7            {c, d}

8            {a, b, c}

9            {a, d, e}

10          {b, d}

Compute the confidence for the association rules  $\{a, e\} \rightarrow \{b\}$ .

0.5

### Question 5

2 / 2 pts

Select the reason why Apriori pruning in the search for frequent itemsets works.

- support count is monotonic with respect to itemsets.
- we search in transaction ID order.
- support count diverges as we add to the itemset.
- support count is anti-monotonic with respect to itemsets.
- it is an excellent heuristic, but it does not work 100% of the time.

### Question 6

2 / 2 pts

Select the true statements about neural network (NN) complexity.

- CNN is always more complex than DNN
  - Adding more hidden layers increases complexity of a NN
  - An Autoencoder is always more complex than a CNN
- Complexity metric defined for a NN may not make sense for other classification engines

**Question 7**

2 / 2 pts

Which of the following is an application of autoencoders?

- 
- feature learning.
- 
- dimensionality reduction.
- 
- Rule mining
- 
- all of the above.

**Question 8**

2 / 2 pts

The number of neurons in the output layer should match the number of classes (Where the number of classes is greater than 2) in a supervised learning task.

- 
- True
- 
- False

**Question 9**

2 / 2 pts

Is an autoencoder the same as performing Principal Component Analysis?

- 
- Not at all
- 
- No but conceptually they can be used for the same purpose
- 
- Autoencoders can do dimensionality reduction, but it is non-linear.

- Yes they have the same purpose.

### Question 10

2 / 2 pts

Increase in size of a convolutional hidden layer would not necessarily increase the performance of a convolutional network.

True

False

Incorrect

### Question 11

0 / 2 pts

Review the following table.

Data	X	Y
x1	1	1
x2	6	6
x3	7	7
x4	3	2
x5	6	8
x6	2	1
x7	2	3

Compute the updated centroids after the first iteration using K-mean clustering with the parameters below.

K:2

Initial centroids: centroid1 = (4,3) and centroid2 = (9,7)

Distance measure: Euclidean

Type your answer in the same format as "centroid1 = (x1,y1) and centroid2 = (x2,y2)"

Each co-ordinate should be in the format A.BC, where A, B and C are integers. If you have 4 as one of the co-ordinates you should write 4.00. Please be mindful of the spaces and other formats.

centroid1 = (2.00,1.75) and centroid2 = (6.33,7.00)

### Question 12

2 / 2 pts

Select all that are not true about DBSCAN.

it has trouble when the clusters have widely varying densities

it is very sensitive to noise and to the size of the data

DBSCAN has a complexity of order  $O(n^3)$

it can be expensive when the computation of nearest neighbors requires computing all pairwise proximities

A and B are correct

B and C are correct

DBSCAN is relatively resistant to noise and can handle clusters of arbitrary shapes and sizes. It has complexity order  $O(n \log n)$

A, C and D are correct

### Question 13

2 / 2 pts

Review the table below.

Cluster	Tennis	Golf	Baseball	Basketball	Football	Hockey	Total	Entropy	Purity
#1	1	1	0	11	4	676	693		
#2	27	89	333	827	253	33	1562		
#3	326	465	8	105	16	29	949		
Total	354	555	341	943	273	738	3204		

Select the collect purity set of the confusion matrix

Cluster #1: 0.98, Cluster #2: 0.53, Cluster #3: 0.49, Total: 0.61

Cluster #1: 0.53, Cluster #2: 0.98, Cluster #3: 0.61, Total: 0.49

Cluster #1: 0.98, Cluster #2: 0.49, Cluster #3: 0.53, Total: 0.61

Cluster #1: 0.53, Cluster #2: 0.49, Cluster #3: 0.61, Total: 0.98

None of above

## Question 14

2 / 2 pts

Select qualities of clusters produced by a good clustering algorithm.

Intra-cluster distances are minimized

Inter-cluster distances are maximized

Number of clusters produced

A and B are correct

good cluster have strong cohesion within the cluster and maximum distance between them.

## Question 15

2 / 2 pts

Review the following table.

Data	X	Y
x1	1	1
x2	6	6
x3	7	7
x4	3	2
x5	6	8
x6	2	1
x7	2	3

Select the correct pair of final centroid and clustering set using K-mean clustering with the parameters below.

K:2

Initial centroids: centroid1 = (4,3) and centroid2 = (9,7)

Distance measure: Euclidean distance



Cluster 1: (x1, x4, x6, x7) with centroid=(2, 1.75), Cluster 2: (x2, x3, x5) with centroid=(6.33, 7)



Cluster 1: (x1, x4, x6, x7) with centroid=(1.75, 2), Cluster 2: (x2, x3, x5) with centroid=(6.33, 7)



Cluster 1: (x1, x4, x6, x7) with centroid=(2, 1.75), Cluster 2: (x2, x3, x5) with centroid=(7, 6.33)



Cluster 1: (x1, x2, x4, x6, x7) with centroid=(2, 1.75), Cluster 2: (x3, x5) with centroid=(6.33, 7)

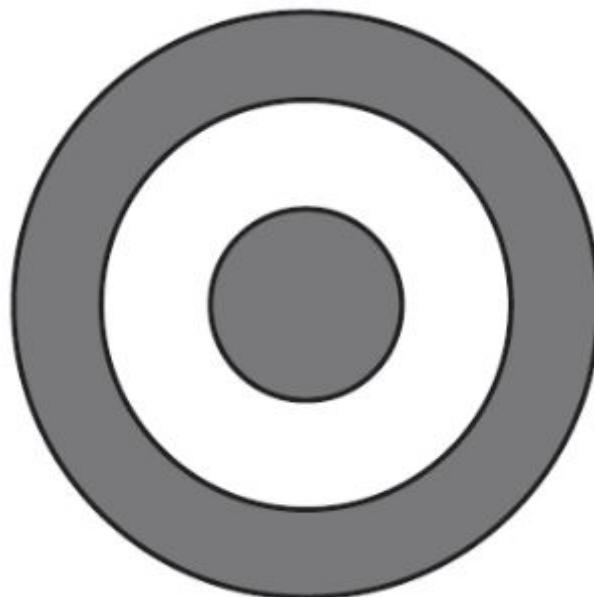


Cluster 1: ( $x_1, x_2, x_4, x_6, x_7$ ) with centroid=(2, 1.75), Cluster 2: ( $x_3, x_5$ ) with centroid=(7, 6.33)

### Question 16

2 / 2 pts

Identify the number of clusters using the center-based, contiguity-based, and density-based definitions for below figure.



center-based: 1 cluster, contiguity-based: 2 clusters, density-based: 2 cluster



center-based: 2 cluster, contiguity-based: 2 clusters, density-based: 1 cluster



center-based: 3 cluster, contiguity-based: 1 clusters, density-based: 3 cluster



center-based: 2 cluster, contiguity-based: 2 clusters, density-based: 2 cluster



None of above

### Question 17

2 / 2 pts

Review the following table.

Data	X	Y
x1	1	1
x2	6	6
x3	7	7
x4	3	2
x5	6	8
x6	2	1
x7	2	3

Compute the sum of squared error (SSE) to evaluate K-mean clustering result with the parameters below.

K:2

Initial centroids: centroid1 = (4,3) and centroid2 = (9,7)

Distance measure: Euclidean distance

7.4167

### Question 18

2 / 2 pts

Review the table below.

Instance	X-Axis	Class
i	-3	-
ii	-2	-
iii	-1	+
iv	0	+
v	1	+
vi	2	-

Select a correct hyper-plane equation of the data below using Non-linear SVM.

## Feature Map ( $\{R^1 \rightarrow R^2\}$ )

---

Y=1

---



---

Y=1.5

---



---

Y=2

---



---

Y=2.5

---



---

None of above

### Question 19

2 / 2 pts

Select the true statement for Nearest Neighbor classification



Although Nearest Neighbor classification has the data dimensionality issue, it can be solved by the scaling technique

k-NN classifier is a typical lazy learner because it spend so much time building the model even if the 'k' is very small

Drawing a circle is a mandatory task to find neighbor classes regardless of the number of 'K'

Determining the optimal value of 'K' is important for k-NN classifier performance.

None of above

### Question 20

2 / 2 pts

Review the table below.

A	B	C	Class
F	F	F	Y
F	F	T	N
F	T	T	N
F	T	T	N
F	F	T	Y
T	F	T	Y
T	F	T	N
T	F	T	N
T	T	T	Y
T	F	T	Y

Select the correct prediction class for a test sample (A=F, B=T, C=F) using the naïve Bayes approach.

'Y'

'N'

Anyone

Quiz Score: **38** out of 40

# Final Exam

**Due** Dec 4 at 11:59pm      **Points** 30      **Questions** 15

**Available** Dec 2 at 12am - Dec 5 at 1am 3 days

**Time Limit** 75 Minutes

## Instructions

No collaboration

75 mins

## Attempt History

	Attempt	Time	Score
LATEST	<a href="#">Attempt 1</a>	26 minutes	30 out of 30

! Correct answers are hidden.

Score for this quiz: **30** out of 30

Submitted Dec 4 at 11:29am

This attempt took 26 minutes.

Question 1		2 / 2 pts
Transaction ID Items Bought		
1	{Milk, Beer, Diapers}	
2	{Bread, Butter, Milk}	
3	{Milk, Diapers, Cookies}	
4	{Bread, Butter, Cookies}	

- 5 {Beer, Cookies, Diapers}
- 6 {Milk, Diapers, Bread, Butter}
- 7 {Bread, Butter, Diapers}
- 8 {Beer, Diapers}
- 9 {Milk, Diapers, Bread, Butter}
- 10 {Beer, Cookies}

What is the maximum size of frequent itemsets that can be extracted  
(assuming minsup > 0.1)?

---

10

---

9

---

4

---

3

## Question 2

2 / 2 pts

Transaction ID Items Bought

- 1 {Milk, Beer, Diapers}
- 2 {Bread, Butter, Milk}

- 3 {Milk, Diapers, Cookies}
- 4 {Bread, Butter, Cookies}
- 5 {Beer, Cookies, Diapers}
- 6 {Milk, Diapers, Bread, Butter}
- 7 {Bread, Butter, Diapers}
- 8 {Beer, Diapers}
- 9 {Milk, Diapers, Bread, Butter}
- 10 {Beer, Cookies}

Find all itemsets (of size 3 or larger) that has the largest support.

{Milk, Beer, Diapers}

{Bread, Butter, Milk}

{Bread, Butter, Diapers}

{Milk, Diapers, Bread}

None of the above.

All of the above.

### Question 3

2 / 2 pts

Transaction ID Items Bought

- 1 {Milk, Beer, Diapers}
- 2 {Bread, Butter, Milk}
- 3 {Milk, Diapers, Cookies}
- 4 {Bread, Butter, Cookies}
- 5 {Beer, Cookies, Diapers}
- 6 {Milk, Diapers, Bread, Butter}
- 7 {Bread, Butter, Diapers}
- 8 {Beer, Diapers}
- 9 {Milk, Diapers, Bread, Butter}
- 10 {Beer, Cookies}

What is the confidence of the rule Beer  $\rightarrow$  Cookies

0.5

#### Question 4

2 / 2 pts

Apriori pruning based on support is a greedy strategy

True

False

**Question 5**

2 / 2 pts

Example of market basket transactions.

Transaction ID Items Bought

1 {a, b, d, e}

2 {b, c, d}

3 {a, b, d, e}

4 {a, c, d, e}

5 {b, c, d, e}

6 {b, d, e}

7 {c, d}

8 {a, b, c}

9 {a, d, e}

10 {b, d}

If the minimum support is set at 40%, how many frequent 3-itemsets will be found?

2

## Question 6

2 / 2 pts

Select the true statements about neural network (NN) complexity.

- 
- CNN is always more complex than DNN
- 
- Adding more hidden layers increases complexity of a NN
- 

- An Autoencoder is always more complex than a CNN
- 



Complexity metric defined for a NN may not make sense for other classification engines

## Question 7

2 / 2 pts

Usage of deep learning is predicated with availability of large amounts of data. Is the data volume the only determining factor to check whether we can use deep learning ?

- 
- Yes, lower volumes lead to overfitting
- 



No, balance among classes in the input data is the only determining factor  
volume does not matter

---



No, balance among classes in the input data is a co-factor along with the  
data volume

---

- None of the above

**Question 8**

2 / 2 pts

The number of neurons in the output layer should match the number of classes (Where the number of classes is greater than 2) in a supervised learning task.

---

True

---

False

**Question 9**

2 / 2 pts

Convolutional networks played an important role in the history of deep learning as indicated by

---

they were the first deep learning models to perform well.

---

they perform better in any type of application.

---

they were used to win many machine classification contests.

---

all of the above

**Question 10**

2 / 2 pts

What is an autoencoder?

---

a neural network that copies its input to its output.

---

a neural network that codes itself.

a neural network that maps an output to an input through a hidden layer.

a neural network that is trained to attempt to copy its input to its output.

### Question 11

2 / 2 pts

Select qualities of clusters produced by a good clustering algorithm.

Intra-cluster distances are minimized

Inter-cluster distances are maximized

Number of clusters produced

A and B are correct

good cluster have strong cohesion within the cluster and maximum distance between them.

### Question 12

2 / 2 pts

Identify the number of clusters using the center-based, contiguity-based, and density-based definitions for below figure.

center-based: 1 cluster, contiguity-based: 2 clusters, density-based: 2 cluster

center-based: 2 cluster, contiguity-based: 2 clusters, density-based: 1 cluster

center-based: 3 cluster, contiguity-based: 1 clusters, density-based: 3 cluster

center-based: 2 cluster, contiguity-based: 2 clusters, density-based: 2 cluster

None of above

### Question 13

2 / 2 pts

Select different aspects of cluster validation.

Determining the clustering tendency of a set of data

To improve the complexity of the algorithm

Determining the 'correct' number of clusters

None of the above

correct these are some of the aspects of cluster validation!

### Question 14

2 / 2 pts

Select all that are not true about DBSCAN.

it has trouble when the clusters have widely varying densities

it is very sensitive to noise and to the size of the data

DBSCAN has a complexity of order  $O(n^3)$

it can be expensive when the computation of nearest neighbors requires computing all pairwise proximities

A and B are correct

B and C are correct

DBSCAN is relatively resistant to noise and can handle clusters of arbitrary shapes and sizes. It has complexity order  $O(n \log n)$

A, C and D are correct

## Question 15

2 / 2 pts

What type of clustering is DBSCAN?

- Center-based
- Contiguous
- Well-separated
- None of above

DBSCAN is a density-based clustering algorithm

Quiz Score: **30** out of 30

# Final Exam

**Due** Dec 4 at 11:59pm      **Points** 30      **Questions** 15

**Available** Dec 2 at 12am - Dec 5 at 1am 3 days      **Time Limit** 75 Minutes

## Instructions

No collaboration

75 mins

## Attempt History

	Attempt	Time	Score
LATEST	<a href="#">Attempt 1</a>	35 minutes	30 out of 30

! Correct answers are hidden.

Score for this quiz: **30** out of 30

Submitted Dec 3 at 6:43pm

This attempt took 35 minutes.

Question 1	2 / 2 pts
Select the reason why Apriori pruning in the search for frequent itemsets works.	
<input type="radio"/> support count is monotonic with respect to itemsets.	
<input type="radio"/> we search in transaction ID order.	
<input type="radio"/> support count diverges as we add to the itemset.	
<input checked="" type="radio"/> support count is anti-monotonic with respect to itemsets.	
<input type="radio"/> it is an excellent heuristic, but it does not work 100% of the time.	

**Question 2****2 / 2 pts**

Transaction ID Items Bought

- 1 {Milk, Beer, Diapers}
- 2 {Bread, Butter, Milk}
- 3 {Milk, Diapers, Cookies}
- 4 {Bread, Butter, Cookies}
- 5 {Beer, Cookies, Diapers}
- 6 {Milk, Diapers, Bread, Butter}
- 7 {Bread, Butter, Diapers}
- 8 {Beer, Diapers}
- 9 {Milk, Diapers, Bread, Butter}
- 10 {Beer, Cookies}

What is the confidence of the rule Beer  $\rightarrow$  Cookies

0.5

**Question 3****2 / 2 pts**

Transaction ID Items Bought

- 1 {Milk, Beer, Diapers}
- 2 {Bread, Butter, Milk}
- 3 {Milk, Diapers, Cookies}
- 4 {Bread, Butter, Cookies}
- 5 {Beer, Cookies, Diapers}
- 6 {Milk, Diapers, Bread, Butter}
- 7 {Bread, Butter, Diapers}
- 8 {Beer, Diapers}
- 9 {Milk, Diapers, Bread, Butter}
- 10 {Beer, Cookies}

Find all itemsets (of size 3 or larger) that has the largest support.

{Milk, Beer, Diapers}

{Bread, Butter, Milk}

{Bread, Butter, Diapers}

{Milk, Diapers, Bread}

None of the above.

All of the above.

#### Question 4

2 / 2 pts

**Transaction ID Items Bought**

- 1 {Milk, Beer, Diapers}
- 2 {Bread, Butter, Milk}
- 3 {Milk, Diapers, Cookies}
- 4 {Bread, Butter, Cookies}
- 5 {Beer, Cookies, Diapers}
- 6 {Milk, Diapers, Bread, Butter}
- 7 {Bread, Butter, Diapers}
- 8 {Beer, Diapers}
- 9 {Milk, Diapers, Bread, Butter}
- 10 {Beer, Cookies}

What is the maximum number of size 3-itemsets that can be derived from this data set.

---

20

---

35

---

40

---

15

---

None of the above.

**Question 5****2 / 2 pts**

Which of the above statements are true for any A, B, and C?

- If  $A \rightarrow B$  then  $B \rightarrow A$ .
- If  $A \rightarrow B$  and  $B \rightarrow C$  then  $A \rightarrow C$ .
- If  $A \rightarrow C$  then  $A \cup B \rightarrow C$ .
- If  $A \cup B \rightarrow C$  then  $A \rightarrow C$ .
- I & II
- I, II, & III
- I, II, & IV
- II & III
- none

**Question 6****2 / 2 pts**

Convolutional networks played an important role in the history of deep learning as indicated by

- they were the first deep learning models to perform well.
- they perform better in any type of application.
- they were used to win many machine classification contests.
- all of the above

**Question 7****2 / 2 pts**

Which of the following is an application of autoencoders?

- 
- feature learning.
- 
- dimensionality reduction.
- 
- Rule mining
- 
- all of the above.

**Question 8****2 / 2 pts**

Increase in size of a convolutional hidden layer would not necessarily increase the performance of a convolutional network.

- 
- True
- 
- False

**Question 9****2 / 2 pts**

What is an autoencoder?

- 
- a neural network that copies its input to its output.
- 
- a neural network that codes itself.
- 
- a neural network that maps an output to an input through a hidden layer.
- 
- a neural network that is trained to attempt to copy its input to its output.

**Question 10****2 / 2 pts**

Select the true statements about neural network (NN) complexity.

- CNN is always more complex than DNN
- Adding more hidden layers increases complexity of a NN

- An Autoencoder is always more complex than a CNN



Complexity metric defined for a NN may not make sense for other classification engines

**Question 11****2 / 2 pts**

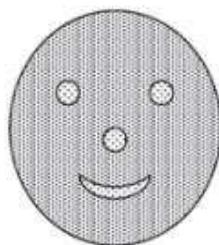
Select qualities of clusters produced by a good clustering algorithm.

- Intra-cluster distances are minimized
- Inter-cluster distances are maximized
- Number of clusters produced
- A and B are correct

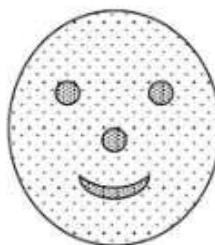
good cluster have strong cohesion within the cluster and maximum distance between them.

**Question 12****2 / 2 pts**

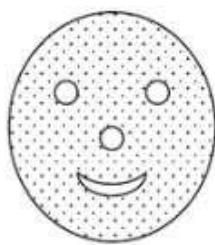
Review the image below.



(a)



(b)



(c)



(d)

**Darkness or number of dots represent density. Lines are used only to distinguish regions and do not represent points.**

If you want to find the patterns represented by the nose, eyes, and mouth using k-Means clustering, select all figures that be well-clustered?

---

 (a)

---

 (b)

---

 (c)

---

 (d)

---

 None of above

### Question 13

2 / 2 pts

Select all statements that are true for Hierarchical clustering.

---

 Although it requires high storage, high computing is not required

---

 Robust to noise

---

 Robust to outliers

It is not easy to decide on the number of cluster because the complicated dendrogram disturbs this decision

 None of above**Question 14**

2 / 2 pts

What is the purpose of cluster analysis?

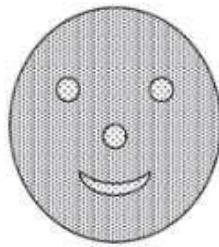
 To avoid finding patterns in noise To compare clustering algorithms To compare two sets of clusters all of the above None of the above

we do cluster validity to avoid clustering noise and find a suitable algorithm for our data

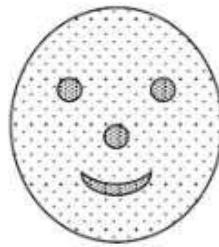
**Question 15**

2 / 2 pts

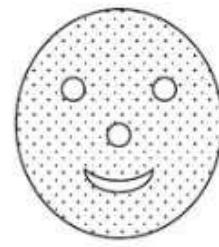
Review the image below.



(a)



(b)



(c)



(d)

**Darkness or number of dots represent density. Lines are used only to distinguish regions and do not represent points.**

If we want to find the patterns represented by the nose, eyes, and mouth using single linkage hierarchical clustering, select all figures that be well-clustered?

---

 (a)

---

 (b)

---

 (c)

---

 (d)

---

 None of above

Quiz Score: **30** out of 30

# Final Exam

Due Dec 4 at 11:59pm      Points 30      Questions 15

Available Dec 2 at 12am - Dec 5 at 1am 3 days

Time Limit 75 Minutes

## Instructions

No collaboration

75 mins

## Attempt History

	Attempt	Time	Score
LATEST	<a href="#">Attempt 1</a>	56 minutes	23 out of 30

! Correct answers are hidden.

Score for this quiz: 23 out of 30

Submitted Dec 3 at 2:56pm

This attempt took 56 minutes.

### Question 1

2 / 2 pts

Transaction ID Items Bought

- 1 {Milk, Beer, Diapers}
- 2 {Bread, Butter, Milk}
- 3 {Milk, Diapers, Cookies}
- 4 {Bread, Butter, Cookies}

- 5 {Beer, Cookies, Diapers}
- 6 {Milk, Diapers, Bread, Butter}
- 7 {Bread, Butter, Diapers}
- 8 {Beer, Diapers}
- 9 {Milk, Diapers, Bread, Butter}
- 10 {Beer, Cookies}

What is the maximum size of frequent itemsets that can be extracted  
(assuming minsup > 0.1)?

---

10

---

9

---

4

---

3

## Question 2

2 / 2 pts

Example of market basket transactions.

Transaction ID Items Bought

- 1 {a, b, d, e}
- 2 {b, c, d}

3             $\{a, b, d, e\}$

4             $\{a, c, d, e\}$

5             $\{b, c, d, e\}$

6             $\{b, d, e\}$

7             $\{c, d\}$

8             $\{a, b, c\}$

9             $\{a, d, e\}$

10           $\{b, d\}$

Compute the confidence for the association rules  $\{a, e\} \rightarrow \{b\}$ .

0.5

### Question 3

2 / 2 pts

Which of the above statements are true for any A, B, and C?

- If  $A \rightarrow B$  then  $B \rightarrow A$ .
- If  $A \rightarrow B$  and  $B \rightarrow C$  then  $A \rightarrow C$ .
- If  $A \rightarrow C$  then  $A \cup B \rightarrow C$ .
- If  $A \cup B \rightarrow C$  then  $A \rightarrow C$ .
- I & II

I, II, & III

I, II, & IV

II & III

none

Incorrect

#### Question 4

0 / 2 pts

Apriori pruning based on support is a greedy strategy

True

False

#### Question 5

2 / 2 pts

Example of market basket transactions.

Transaction ID Items Bought

1 {a, b, d, e}

2 {b, c, d}

3 {a, b, d, e}

4 {a, c, d, e}

5            {b, c, d, e}

6            {b, d, e}

7            {c, d}

8            {a, b, c}

9            {a, d, e}

10          {b, d}

Compute the support for itemset {b} -> {c,d}.

0.2

## Question 6

2 / 2 pts

What is an autoencoder?

a neural network that copies its input to its output.

a neural network that codes itself.

a neural network that maps an output to an input through a hidden layer.

a neural network that is trained to attempt to copy its input to its output.

Incorrect

## Question 7

0 / 2 pts

The number of neurons in the output layer should match the number of classes (Where the number of classes is greater than 2) in a supervised learning task.

---

True

---

False

### Question 8

2 / 2 pts

Select all that is true about Restricted Boltzman Machine (RBM)

---

Can be used for recommendation systems

---

RBM training is often probabilistic

---

RBM is a kind of Deep Belief Network

---

None of the above.

Incorrect

### Question 9

0 / 2 pts

Increase in size of a convolutional hidden layer would not necessarily increase the performance of a convolutional network.

---

True

---

False

## Question 10

2 / 2 pts

Select the true statements about neural network (NN) complexity.

- CNN is always more complex than DNN
- Adding more hidden layers increases complexity of a NN

- An Autoencoder is always more complex than a CNN



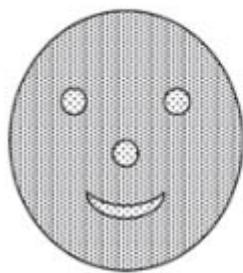
Complexity metric defined for a NN may not make sense for other classification engines

Partial

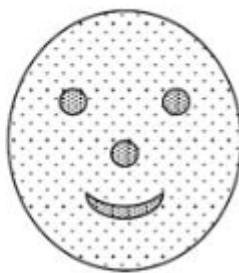
## Question 11

1 / 2 pts

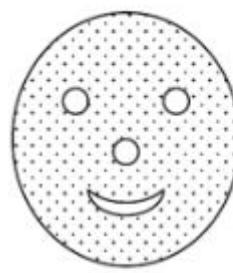
Review the image below.



(a)



(b)



(c)



(d)

**Darkness or number of dots represent density. Lines are used only to distinguish regions and do not represent points.**

If you want to find the patterns represented by the nose, eyes, and mouth using k-Means clustering, select all figures that be well-clustered?

- (a)

(b)

(c)

(d)

None of above

### Question 12

2 / 2 pts

What type of clustering is DBSCAN?

Center-based

Contiguous

Well-separated

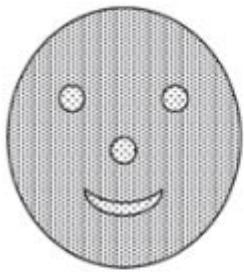
None of above

DBSCAN is a density-based clustering algorithm

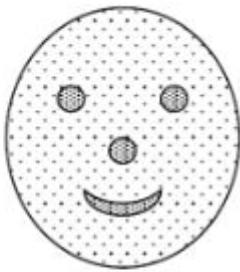
### Question 13

2 / 2 pts

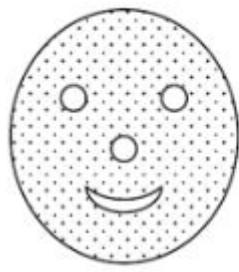
Review the image below.



(a)



(b)



(c)



(d)

**Darkness or number of dots represent density. Lines are used only to distinguish regions and do not represent points.**

If we want to find the patterns represented by the nose, eyes, and mouth using single linkage hierarchical clustering, select all figures that be well-clustered?

(a)

(b)

(c)

(d)

None of above

### Question 14

2 / 2 pts

Select all that are not true about DBSCAN.

it has trouble when the clusters have widely varying densities

it is very sensitive to noise and to the size of the data

DBSCAN has a complexity of order  $O(n^3)$



it can be expensive when the computation of nearest neighbors requires computing all pairwise proximities

- 
- A and B are correct
- 

- B and C are correct

DBSCAN is relatively resistant to noise and can handle clusters of arbitrary shapes and sizes. it has complexity order  $O(n \log n)$

- 
- A, C and D are correct
- 

### Question 15

2 / 2 pts

Select qualities of clusters produced by a good clustering algorithm.

- 
- Intra-cluster distances are minimized
- 

- Inter-cluster distances are maximized
- 

- Number of clusters produced
- 

- A and B are correct

good cluster have strong cohesion within the cluster and maximum distance between them.

Quiz Score: 23 out of 30

# Final Exam

**Due** Dec 9 at 1:15am      **Points** 40      **Questions** 20

**Available** Dec 6 at 12am - Dec 9 at 2am 3 days

**Time Limit** 75 Minutes

## Instructions

No collaboration

75 mins

## Attempt History

	Attempt	Time	Score
LATEST	<a href="#"><u>Attempt 1</u></a>	39 minutes	40 out of 40

❗ Correct answers are hidden.

Score for this quiz: **40** out of 40

Submitted Dec 8 at 5:10pm

This attempt took 39 minutes.

Question 1	2 / 2 pts
Select the reason why Apriori pruning in the search for frequent itemsets works.	
<input type="radio"/> support count is monotonic with respect to itemsets.	
<input type="radio"/> we search in transaction ID order.	
<input type="radio"/> support count diverges as we add to the itemset.	
<input checked="" type="radio"/> support count is anti-monotonic with respect to itemsets.	

- it is an excellent heuristic, but it does not work 100% of the time.

**Question 2****2 / 2 pts**

Example of market basket transactions.

Transaction ID Items Bought

1 {a, b, d, e}

2 {b, c, d}

3 {a, b, d, e}

4 {a, c, d, e}

5 {b, c, d, e}

6 {b, d, e}

7 {c, d}

8 {a, b, c}

9 {a, d, e}

10 {b, d}

Compute the support for itemset {b,c,d}.

0.2

**Question 3**

**2 / 2 pts**

## Transaction ID Items Bought

- 1 {Milk, Beer, Diapers}
- 2 {Bread, Butter, Milk}
- 3 {Milk, Diapers, Cookies}
- 4 {Bread, Butter, Cookies}
- 5 {Beer, Cookies, Diapers}
- 6 {Milk, Diapers, Bread, Butter}
- 7 {Bread, Butter, Diapers}
- 8 {Beer, Diapers}
- 9 {Milk, Diapers, Bread, Butter}
- 10 {Beer, Cookies}

What is the confidence of the rule Beer  $\rightarrow$  Cookies

0.5

**Question 4****2 / 2 pts**

Example of market basket transactions.

## Transaction ID Items Bought

1 {a, b, d, e}

2 {b, c, d}

3 {a, b, d, e}

4 {a, c, d, e}

5 {b, c, d, e}

6 {b, d, e}

7 {c, d}

8 {a, b, c}

9 {a, d, e}

10 {b, d}

If the minimum support is set at 40%, how many frequent 3-itemsets will be found? Note that if a 3 itemset is a subset of a larger itemset, it counts as one occurrence.

2

**Question 5****2 / 2 pts**

Apriori pruning based on support is a greedy strategy but is not optimal

True False**Question 6****2 / 2 pts**

Usage of deep learning is predicated with availability of large amounts of data. Is the data volume the only determining factor to check whether we can use deep learning ?

 Yes, lower volumes lead to overfitting

No, balance among classes in the input data is the only determining factor volume does not matter

No, balance among classes in the input data is a co-factor along with the data volume

 None of the above**Question 7****2 / 2 pts**

The number of neurons in the output layer should match the number of classes (Where the number of classes is greater than 2) in a supervised learning task.

 True False

**Question 8****2 / 2 pts**

Is an autoencoder the same as performing Principal Component Analysis?

- 
- Not at all
- 
- No but conceptually they can be used for the same purpose
- 
- Autoencoders can do dimensionality reduction, but it is non-linear.
- 
- Yes they have the same purpose.

**Question 9****2 / 2 pts**

Consider a movie recommendation application, where an user can log in and select movies they like based on their previous preferences.

Which neural network architecture would be suitable to complete this task

- 
- Fully-Connected Neural Network.
- 
- Convolutional Neural Network.
- 
- Recurrent Neural Network.
- 
- Restricted Boltzmann Machine.

**Question 10****2 / 2 pts**

Increase in size of a convolutional hidden layer would not necessarily increase the performance of a convolutional network.

True

False

### Question 11

2 / 2 pts

Review the table below.

Cluster	Tennis	Golf	Baseball	Basketball	Football	Hockey	Total	Entropy	Purity
#1	1	1	0	11	4	676	693		
#2	27	89	333	827	253	33	1562		
#3	326	465	8	105	16	29	949		
Total	354	555	341	943	273	738	3204		

Select the collect purity set of the confusion matrix

Cluster #1: 0.98, Cluster #2: 0.53, Cluster #3: 0.49, Total: 0.61

Cluster #1: 0.53, Cluster #2: 0.98, Cluster #3: 0.61, Total: 0.49

Cluster #1: 0.98, Cluster #2: 0.49, Cluster #3: 0.53, Total: 0.61

Cluster #1: 0.53, Cluster #2: 0.49, Cluster #3: 0.61, Total: 0.98

None of above

### Question 12

2 / 2 pts

Review the following table.

Data	X	Y
x1	1	1
x2	6	6
x3	7	7
x4	3	2
x5	6	8
x6	2	1
x7	2	3

Select the correct pair of final centroid and clustering set using K-mean clustering with the parameters below.

K:2

Initial centroids: centroid1 = (4,3) and centroid2 = (9,7)

Distance measure: Euclidean distance



Cluster 1: (x1, x4, x6, x7) with centroid=(2, 1.75), Cluster 2: (x2, x3, x5) with centroid=(6.33, 7)



Cluster 1: (x1, x4, x6, x7) with centroid=(1.75, 2), Cluster 2: (x2, x3, x5) with centroid=(6.33, 7)



Cluster 1: (x1, x4, x6, x7) with centroid=(2, 1.75), Cluster 2: (x2, x3, x5) with centroid=(7, 6.33)



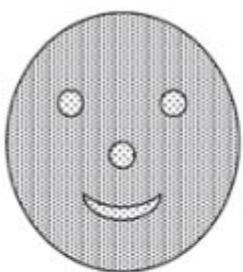
Cluster 1: (x1, x2, x4, x6, x7) with centroid=(2, 1.75), Cluster 2: (x3, x5) with centroid=(6.33, 7)



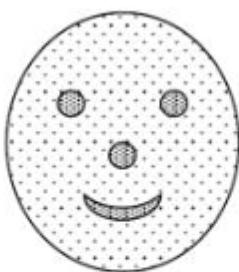
Cluster 1: (x1, x2, x4, x6, x7) with centroid=(2, 1.75), Cluster 2: (x3, x5) with centroid=(7, 6.33)

**Question 13****2 / 2 pts**

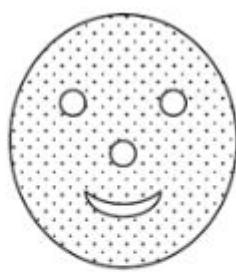
Review the image below.



(a)



(b)



(c)



(d)

**Darkness or number of dots represent density. Lines are used only to distinguish regions and do not represent points.**

If you want to find the patterns represented by the nose, eyes, and mouth using k-Means clustering, select all figures that be well-clustered?

 (a) (b) (c) (d) None of above**Question 14****2 / 2 pts**

Select qualities of clusters produced by a good clustering algorithm.

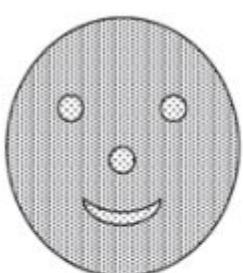
- Intra-cluster distances are minimized
- Inter-cluster distances are maximized
- Number of clusters produced
- A and B are correct

good cluster have strong cohesion within the cluster and maximum distance between them.

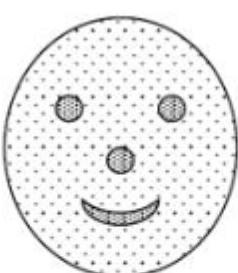
### Question 15

2 / 2 pts

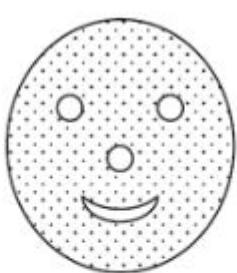
Review the image below.



(a)



(b)



(c)



(d)

**Darkness or number of dots represent density. Lines are used only to distinguish regions and do not represent points.**

If we want to find the patterns represented by the nose, eyes, and mouth using single linkage hierarchical clustering, select all figures that be well-clustered?

(a)

(b)

(c)

(d) None of above**Question 16****2 / 2 pts**

Select all different type of clusters.

 Well-separated Center-based Contiguous Density-based All of the above

all are type of clustering

**Question 17****2 / 2 pts**

Review the following table.

	A	B	C	D	E
A	0	9	3	6	11
B	9	0	7	5	10
C	3	7	0	9	2
D	6	5	9	0	8
E	11	10	2	8	0

Select the correct DBSCAN clustering result with the parameters below

epsilon = 5, Minpts = 2

- cluster1: {A, C, E}, cluster2: {B,D}
- cluster1: {A}, cluster2: {B}, cluster3: {D}, and cluster4: {C, E}
- cluster1: {A, B}, cluster2: {C}, and cluster3: {D}, cluster4: {E}
- cluster1: {A}, cluster2: {B, E}, cluster3: {C}, and cluster4: {D}
- None of above

### Question 18

2 / 2 pts

Review the table below.

A	B	C	Class
F	F	F	Y
F	F	T	N
F	T	T	N
F	T	T	N
F	F	T	Y
T	F	T	Y
T	F	T	N
T	F	T	N
T	T	T	Y
T	F	T	Y

Select the correct prediction class for a test sample (A=F, B=T, C=F) using the naïve Bayes approach.

'Y'

'N'

Anyone

### Question 19

2 / 2 pts

Select the true statement for Nearest Neighbor classification

Although Nearest Neighbor classification has the data dimensionality issue, it can be solved by the scaling technique

k-NN classifier is a typical lazy learner because it spend so much time building the model even if the 'k' is very small

Drawing a circle is a mandatory task to find neighbor classes regardless of the number of 'K'

Determining the optimal value of 'K' is important for k-NN classifier performance.

None of above

### Question 20

2 / 2 pts

Review the table below.

Instance	X-Axis	Class
i	-3	-
ii	-2	-
iii	-1	+
iv	0	+
v	1	+
vi	2	-

Select all the support vector instances using Non-linear SVM

Feature Map ( $\{R^1 \rightarrow R^2\}$ )

 i

<input checked="" type="checkbox"/> ii
<input checked="" type="checkbox"/> iii
<input type="checkbox"/> iv
<input checked="" type="checkbox"/> v
<input checked="" type="checkbox"/> vi

Quiz Score: **40** out of 40

# Final Exam

**Due** Dec 8 at 1:15am      **Points** 40      **Questions** 20

**Available** Dec 6 at 12am - Dec 8 at 2am 2 days

**Time Limit** 75 Minutes

## Instructions

No collaboration

75 mins

## Attempt History

	<b>Attempt</b>	<b>Time</b>	<b>Score</b>
<b>LATEST</b>	<a href="#"><u>Attempt 1</u></a>	50 minutes	38 out of 40

! Correct answers are hidden.

Score for this quiz: **38** out of 40

Submitted Dec 6 at 2:37pm

This attempt took 50 minutes.

<b>Question 1</b>		<b>2 / 2 pts</b>
Example of market basket transactions.		
Transaction ID Items Bought		
1	{a, b, d, e}	
2	{b, c, d}	
3	{a, b, d, e}	

- 4      {a, c, d, e}
- 5      {b, c, d, e}
- 6      {b, d, e}
- 7      {c, d}
- 8      {a, b, c}
- 9      {a, d, e}
- 10     {b, d}

Compute the support for itemset {b,c,d}.

0.2

## Question 2

2 / 2 pts

Example of market basket transactions.

### Transaction ID Items Bought

- | Transaction ID | Items Bought |
|----------------|--------------|
| 1              | {a, b, d, e} |
| 2              | {b, c, d}    |
| 3              | {a, b, d, e} |
| 4              | {a, c, d, e} |

5      {b, c, d, e}

6      {b, d, e}

7      {c, d}

8      {a, b, c}

9      {a, d, e}

10     {b, d}

Compute the confidence for the association rules  $\{a, e\} \rightarrow \{b\}$ .

0.5

### Question 3

2 / 2 pts

Transaction ID Items Bought

1      {Milk, Beer, Diapers}

2      {Bread, Butter, Milk}

3      {Milk, Diapers, Cookies}

4      {Bread, Butter, Cookies}

5      {Beer, Cookies, Diapers}

6      {Milk, Diapers, Bread, Butter}

- 7      {Bread, Butter, Diapers}
- 8      {Beer, Diapers}
- 9      {Milk, Diapers, Bread, Butter}
- 10     {Beer, Cookies}

What is the confidence of the rule Beer  $\rightarrow$  Cookies

0.5

#### Question 4

2 / 2 pts

Example of market basket transactions.

#### Transaction ID Items Bought

- 1      {a, b, d, e}
- 2      {b, c, d}
- 3      {a, b, d, e}
- 4      {a, c, d, e}
- 5      {b, c, d, e}
- 6      {b, d, e}
- 7      {c, d}

8            {a, b, c}

9            {a, d, e}

10          {b, d}

If the minimum support is set at 40%, how many frequent 3-itemsets will be found? Note that if a 3 itemset is a subset of a larger itemset, it counts as one occurrence.

2

### Question 5

2 / 2 pts

Transaction ID Items Bought

1            {Milk, Beer, Diapers}

2            {Bread, Butter, Milk}

3            {Milk, Diapers, Cookies}

4            {Bread, Butter, Cookies}

5            {Beer, Cookies, Diapers}

6            {Milk, Diapers, Bread, Butter}

7            {Bread, Butter, Diapers}

8            {Beer, Diapers}

9 {Milk, Diapers, Bread, Butter}

10 {Beer, Cookies}

What is the maximum number of size 3-itemsets that can be derived from this data set. Note that the itemsets derived may not be in the dataset.

20

35

40

15

None of the above.

### Question 6

2 / 2 pts

Which of the following is an application of autoencoders?

feature learning.

dimensionality reduction.

Rule mining

all of the above.

### Question 7

2 / 2 pts

The number of neurons in the output layer should match the number of classes (Where the number of classes is greater than 2) in a supervised learning task.

---

True

---

False

**Question 8****2 / 2 pts**

What is an autoencoder?

---

a neural network that copies its input to its output.

---

a neural network that codes itself.

---

a neural network that maps an output to an input through a hidden layer.

---

a neural network that is trained to attempt to copy its input to its output.

**Question 9****2 / 2 pts**

Consider a movie recommendation application, where an user can log in and select movies they like based on their previous preferences.

Which neural network architecture would be suitable to complete this task

---

Fully-Connected Neural Network.

---

Convolutional Neural Network.

---

- Recurrent Neural Network.
- 
- Restricted Boltzmann Machine.

Incorrect

**Question 10****0 / 2 pts**

Increase in size of a convolutional hidden layer would not necessarily increase the performance of a convolutional network.

- 
- True
- 
- False

**Question 11****2 / 2 pts**

Select different aspects of cluster validation.

- 
- Determining the clustering tendency of a set of data
- 
- To improve the complexity of the algorithm
- 
- Determining the 'correct' number of clusters
- 
- None of the above

correct these are some of the aspects of cluster validation!

**Question 12****2 / 2 pts**

What is the purpose of cluster analysis?

- To avoid finding patterns in noise
- To compare clustering algorithms
- To compare two sets of clusters
- all of the above
- None of the above

we do cluster validity to avoid clustering noise and find a suitable algorithm for our data

**Question 13****2 / 2 pts**

Select qualities of clusters produced by a good clustering algorithm.

- Intra-cluster distances are minimized
- Inter-cluster distances are maximized
- Number of clusters produced
- A and B are correct

good cluster have strong cohesion within the cluster and maximum distance between them.

**Question 14****2 / 2 pts**

Select all statements that are true for Hierarchical clustering.

Although it requires high storage, high computing is not required

Robust to noise

Robust to outliers

It is not easy to decide on the number of cluster because the complicated dendrogram disturbs this decision

None of above

**Question 15****2 / 2 pts**

Review the following table.

Data	X	Y
A	1	2
B	2	2
C	2	3
D	8	7
E	8	8
F	8	2

Select the correct number of core points and DBSCAN clustering result with the parameters below.

Eps: 5

Minimum samples (MinPts):3

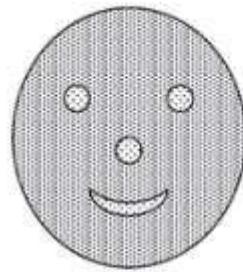
### Distance measurement: Euclidean distance

- 5, cluster1: {A,B,C}, cluster2: {D,E,F}
- 6, cluster1: {A,B}, cluster2: {C,D,E}, cluster3: {F}
- 4, cluster1: {A,B,C}, cluster2: {D, E, F}
- 5, cluster1: {A}, cluster2: {B}, cluster3: {C}, cluster4: {D}, cluster5: {E}
- None of above

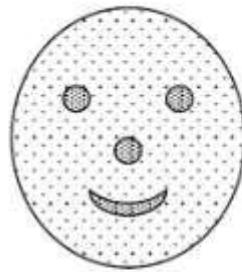
### Question 16

2 / 2 pts

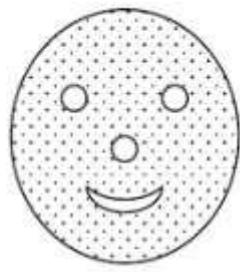
Review the image below.



(a)



(b)



(c)



(d)

**Darkness or number of dots represent density. Lines are used only to distinguish regions and do not represent points.**

If you want to find the patterns represented by the nose, eyes, and mouth using k-Means clustering, select all figures that be well-clustered?

 (a) (b) (c)

(d) None of above**Question 17****2 / 2 pts**

Select all that are not true about DBSCAN.

- it has trouble when the clusters have widely varying densities
- it is very sensitive to noise and to the size of the data
- DBSCAN has a complexity of order  $O(n^3)$
- 
- it can be expensive when the computation of nearest neighbors requires computing all pairwise proximities
- A and B are correct
- B and C are correct

DBSCAN is relatively resistant to noise and can handle clusters of arbitrary shapes and sizes. It has complexity order  $O(n \log n)$

- A, C and D are correct

**Question 18****2 / 2 pts**

Review the table below.

Instance	X-Axis	Class
i	-3	-
ii	-2	-
iii	-1	+
iv	0	+
v	1	+
vi	2	-

Select a correct hyper-plane equation of the data below using Non-linear SVM.

## Feature Map ( $\{R^1 \rightarrow R^2\}$ )

Y=1

Y=1.5

Y=2

Y=2.5

None of above

### Question 19

2 / 2 pts

Review the table below.

Instance	X-Axis	Class
i	-3	-
ii	-2	-
iii	-1	+
iv	0	+
v	1	+
vi	2	-

Select all the support vector instances using Non-linear SVM

## Feature Map ( $\{R^1 \rightarrow R^2\}$ )

i

ii

iii

iv

v

vi

### Question 20

2 / 2 pts

Review the table below.

A	B	C	Class
F	F	F	Y
F	F	T	N
F	T	T	N
F	T	T	N
F	F	T	Y
T	F	T	Y
T	F	T	N
T	F	T	N
T	T	T	Y
T	F	T	Y

Select the correct prediction class for a test sample (A=F, B=T, C=F) using the naïve Bayes approach.

'Y'

'N'

Anyone

Quiz Score: **38** out of 40

# Final Exam

**Due** Dec 4 at 11:59pm      **Points** 30      **Questions** 15

**Available** Dec 2 at 12am - Dec 5 at 1am 3 days      **Time Limit** 75 Minutes

## Instructions

No collaboration

75 mins

## Attempt History

	Attempt	Time	Score
LATEST	<a href="#">Attempt 1</a>	35 minutes	30 out of 30

! Correct answers are hidden.

Score for this quiz: **30** out of 30

Submitted Dec 3 at 6:43pm

This attempt took 35 minutes.

Question 1	2 / 2 pts
Select the reason why Apriori pruning in the search for frequent itemsets works.	
<input type="radio"/> support count is monotonic with respect to itemsets.	
<input type="radio"/> we search in transaction ID order.	
<input type="radio"/> support count diverges as we add to the itemset.	
<input checked="" type="radio"/> support count is anti-monotonic with respect to itemsets.	
<input type="radio"/> it is an excellent heuristic, but it does not work 100% of the time.	

**Question 2****2 / 2 pts**

Transaction ID Items Bought

- 1 {Milk, Beer, Diapers}
- 2 {Bread, Butter, Milk}
- 3 {Milk, Diapers, Cookies}
- 4 {Bread, Butter, Cookies}
- 5 {Beer, Cookies, Diapers}
- 6 {Milk, Diapers, Bread, Butter}
- 7 {Bread, Butter, Diapers}
- 8 {Beer, Diapers}
- 9 {Milk, Diapers, Bread, Butter}
- 10 {Beer, Cookies}

What is the confidence of the rule Beer  $\rightarrow$  Cookies

0.5

**Question 3****2 / 2 pts**

Transaction ID Items Bought

- 1 {Milk, Beer, Diapers}
- 2 {Bread, Butter, Milk}
- 3 {Milk, Diapers, Cookies}
- 4 {Bread, Butter, Cookies}
- 5 {Beer, Cookies, Diapers}
- 6 {Milk, Diapers, Bread, Butter}
- 7 {Bread, Butter, Diapers}
- 8 {Beer, Diapers}
- 9 {Milk, Diapers, Bread, Butter}
- 10 {Beer, Cookies}

Find all itemsets (of size 3 or larger) that has the largest support.

{Milk, Beer, Diapers}

{Bread, Butter, Milk}

{Bread, Butter, Diapers}

{Milk, Diapers, Bread}

None of the above.

All of the above.

#### Question 4

2 / 2 pts

**Transaction ID Items Bought**

- 1 {Milk, Beer, Diapers}
- 2 {Bread, Butter, Milk}
- 3 {Milk, Diapers, Cookies}
- 4 {Bread, Butter, Cookies}
- 5 {Beer, Cookies, Diapers}
- 6 {Milk, Diapers, Bread, Butter}
- 7 {Bread, Butter, Diapers}
- 8 {Beer, Diapers}
- 9 {Milk, Diapers, Bread, Butter}
- 10 {Beer, Cookies}

What is the maximum number of size 3-itemsets that can be derived from this data set.

---

20

---

35

---

40

---

15

---

None of the above.

**Question 5****2 / 2 pts**

Which of the above statements are true for any A, B, and C?

- If  $A \rightarrow B$  then  $B \rightarrow A$ .
- If  $A \rightarrow B$  and  $B \rightarrow C$  then  $A \rightarrow C$ .
- If  $A \rightarrow C$  then  $A \cup B \rightarrow C$ .
- If  $A \cup B \rightarrow C$  then  $A \rightarrow C$ .
- I & II
- I, II, & III
- I, II, & IV
- II & III
- none

**Question 6****2 / 2 pts**

Convolutional networks played an important role in the history of deep learning as indicated by

- they were the first deep learning models to perform well.
- they perform better in any type of application.
- they were used to win many machine classification contests.
- all of the above

**Question 7****2 / 2 pts**

Which of the following is an application of autoencoders?

---

feature learning.

---

dimensionality reduction.

---

Rule mining

---

all of the above.

---

**Question 8****2 / 2 pts**

Increase in size of a convolutional hidden layer would not necessarily increase the performance of a convolutional network.

---

True

---

False

---

**Question 9****2 / 2 pts**

What is an autoencoder?

---

a neural network that copies its input to its output.

---

a neural network that codes itself.

---

a neural network that maps an output to an input through a hidden layer.

---

a neural network that is trained to attempt to copy its input to its output.

---

**Question 10****2 / 2 pts**

Select the true statements about neural network (NN) complexity.

- CNN is always more complex than DNN
- Adding more hidden layers increases complexity of a NN

- An Autoencoder is always more complex than a CNN



Complexity metric defined for a NN may not make sense for other classification engines

**Question 11****2 / 2 pts**

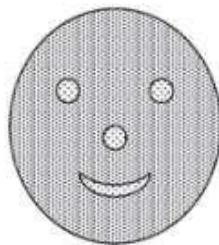
Select qualities of clusters produced by a good clustering algorithm.

- Intra-cluster distances are minimized
- Inter-cluster distances are maximized
- Number of clusters produced
- A and B are correct

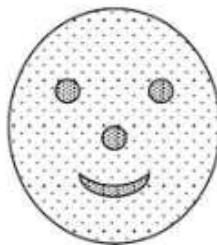
good cluster have strong cohesion within the cluster and maximum distance between them.

**Question 12****2 / 2 pts**

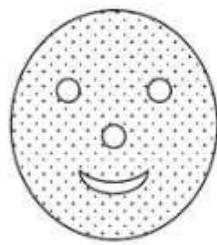
Review the image below.



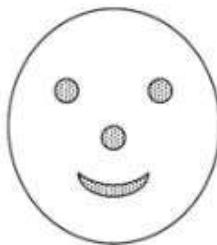
(a)



(b)



(c)



(d)

**Darkness or number of dots represent density. Lines are used only to distinguish regions and do not represent points.**

If you want to find the patterns represented by the nose, eyes, and mouth using k-Means clustering, select all figures that be well-clustered?

---

 (a)

---

 (b)

---

 (c)

---

 (d)

---

 None of above

### Question 13

2 / 2 pts

Select all statements that are true for Hierarchical clustering.

---

 Although it requires high storage, high computing is not required

---

 Robust to noise

---

 Robust to outliers

It is not easy to decide on the number of cluster because the complicated dendrogram disturbs this decision

 None of above**Question 14**

2 / 2 pts

What is the purpose of cluster analysis?

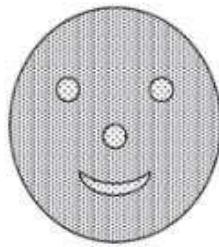
 To avoid finding patterns in noise To compare clustering algorithms To compare two sets of clusters all of the above None of the above

we do cluster validity to avoid clustering noise and find a suitable algorithm for our data

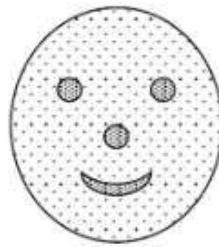
**Question 15**

2 / 2 pts

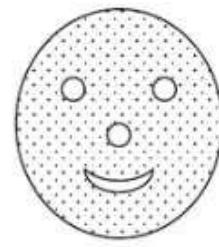
Review the image below.



(a)



(b)



(c)



(d)

**Darkness or number of dots represent density. Lines are used only to distinguish regions and do not represent points.**

If we want to find the patterns represented by the nose, eyes, and mouth using single linkage hierarchical clustering, select all figures that be well-clustered?

---

(a)

---

(b)

---

(c)

---

(d)

---

None of above

---

Quiz Score: **30** out of 30

# Final Exam

Due Dec 4 at 11:59pm      Points 30      Questions 15

Available Dec 2 at 12am - Dec 5 at 1am 3 days

Time Limit 75 Minutes

## Instructions

No collaboration

75 mins

## Attempt History

	Attempt	Time	Score
LATEST	<a href="#">Attempt 1</a>	56 minutes	23 out of 30

! Correct answers are hidden.

Score for this quiz: 23 out of 30

Submitted Dec 3 at 2:56pm

This attempt took 56 minutes.

### Question 1

2 / 2 pts

Transaction ID Items Bought

- 1 {Milk, Beer, Diapers}
- 2 {Bread, Butter, Milk}
- 3 {Milk, Diapers, Cookies}
- 4 {Bread, Butter, Cookies}

- 5 {Beer, Cookies, Diapers}
- 6 {Milk, Diapers, Bread, Butter}
- 7 {Bread, Butter, Diapers}
- 8 {Beer, Diapers}
- 9 {Milk, Diapers, Bread, Butter}
- 10 {Beer, Cookies}

What is the maximum size of frequent itemsets that can be extracted  
(assuming minsup > 0.1)?

---

10

---

9

---

4

---

3

## Question 2

2 / 2 pts

Example of market basket transactions.

Transaction ID Items Bought

- 1 {a, b, d, e}
- 2 {b, c, d}

3             $\{a, b, d, e\}$

4             $\{a, c, d, e\}$

5             $\{b, c, d, e\}$

6             $\{b, d, e\}$

7             $\{c, d\}$

8             $\{a, b, c\}$

9             $\{a, d, e\}$

10           $\{b, d\}$

Compute the confidence for the association rules  $\{a, e\} \rightarrow \{b\}$ .

0.5

### Question 3

2 / 2 pts

Which of the above statements are true for any A, B, and C?

- If  $A \rightarrow B$  then  $B \rightarrow A$ .
- If  $A \rightarrow B$  and  $B \rightarrow C$  then  $A \rightarrow C$ .
- If  $A \rightarrow C$  then  $A \cup B \rightarrow C$ .
- If  $A \cup B \rightarrow C$  then  $A \rightarrow C$ .
- I & II

I, II, & III

I, II, & IV

II & III

none

Incorrect

#### Question 4

0 / 2 pts

Apriori pruning based on support is a greedy strategy

True

False

#### Question 5

2 / 2 pts

Example of market basket transactions.

Transaction ID Items Bought

1 {a, b, d, e}

2 {b, c, d}

3 {a, b, d, e}

4 {a, c, d, e}

5            {b, c, d, e}

6            {b, d, e}

7            {c, d}

8            {a, b, c}

9            {a, d, e}

10          {b, d}

Compute the support for itemset {b} -> {c,d}.

0.2

## Question 6

2 / 2 pts

What is an autoencoder?

a neural network that copies its input to its output.

a neural network that codes itself.

a neural network that maps an output to an input through a hidden layer.

a neural network that is trained to attempt to copy its input to its output.

Incorrect

## Question 7

0 / 2 pts

The number of neurons in the output layer should match the number of classes (Where the number of classes is greater than 2) in a supervised learning task.

---

True

---

False

### Question 8

2 / 2 pts

Select all that is true about Restricted Boltzman Machine (RBM)

---

Can be used for recommendation systems

---

RBM training is often probabilistic

---

RBM is a kind of Deep Belief Network

---

None of the above.

Incorrect

### Question 9

0 / 2 pts

Increase in size of a convolutional hidden layer would not necessarily increase the performance of a convolutional network.

---

True

---

False

## Question 10

2 / 2 pts

Select the true statements about neural network (NN) complexity.

- CNN is always more complex than DNN
- Adding more hidden layers increases complexity of a NN

- An Autoencoder is always more complex than a CNN



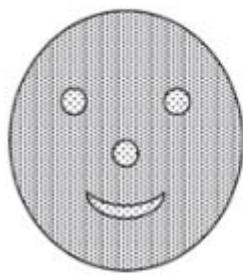
Complexity metric defined for a NN may not make sense for other classification engines

Partial

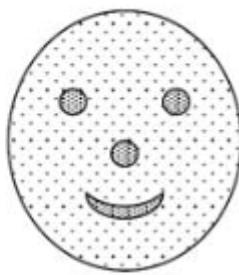
## Question 11

1 / 2 pts

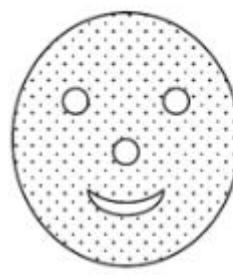
Review the image below.



(a)



(b)



(c)



(d)

**Darkness or number of dots represent density. Lines are used only to distinguish regions and do not represent points.**

If you want to find the patterns represented by the nose, eyes, and mouth using k-Means clustering, select all figures that be well-clustered?

- (a)

(b)

(c)

(d)

None of above

### Question 12

2 / 2 pts

What type of clustering is DBSCAN?

Center-based

Contiguous

Well-separated

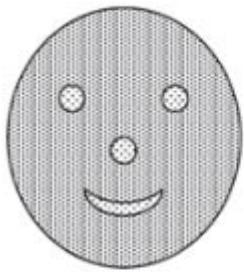
None of above

DBSCAN is a density-based clustering algorithm

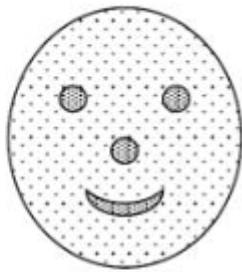
### Question 13

2 / 2 pts

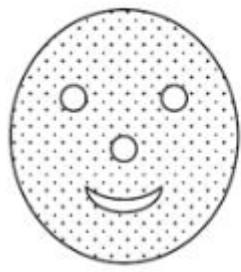
Review the image below.



(a)



(b)



(c)



(d)

**Darkness or number of dots represent density. Lines are used only to distinguish regions and do not represent points.**

If we want to find the patterns represented by the nose, eyes, and mouth using single linkage hierarchical clustering, select all figures that be well-clustered?

(a)

(b)

(c)

(d)

None of above

### Question 14

2 / 2 pts

Select all that are not true about DBSCAN.

it has trouble when the clusters have widely varying densities

it is very sensitive to noise and to the size of the data

DBSCAN has a complexity of order  $O(n^3)$



it can be expensive when the computation of nearest neighbors requires computing all pairwise proximities

A and B are correct

B and C are correct

DBSCAN is relatively resistant to noise and can handle clusters of arbitrary shapes and sizes. it has complexity order  $O(n \log n)$

A, C and D are correct

### Question 15

2 / 2 pts

Select qualities of clusters produced by a good clustering algorithm.

Intra-cluster distances are minimized

Inter-cluster distances are maximized

Number of clusters produced

A and B are correct

good cluster have strong cohesion within the cluster and maximum distance between them.

Quiz Score: 23 out of 30

# Final Exam

**Due** Dec 4 at 11:59pm      **Points** 30      **Questions** 15

**Available** Dec 2 at 12am - Dec 5 at 1am 3 days

**Time Limit** 75 Minutes

## Instructions

No collaboration

75 mins

## Attempt History

	Attempt	Time	Score
LATEST	<a href="#">Attempt 1</a>	26 minutes	30 out of 30

! Correct answers are hidden.

Score for this quiz: 30 out of 30

Submitted Dec 4 at 11:29am

This attempt took 26 minutes.

Question 1		2 / 2 pts
Transaction ID Items Bought		
1	{Milk, Beer, Diapers}	
2	{Bread, Butter, Milk}	
3	{Milk, Diapers, Cookies}	
4	{Bread, Butter, Cookies}	

- 5 {Beer, Cookies, Diapers}
- 6 {Milk, Diapers, Bread, Butter}
- 7 {Bread, Butter, Diapers}
- 8 {Beer, Diapers}
- 9 {Milk, Diapers, Bread, Butter}
- 10 {Beer, Cookies}

What is the maximum size of frequent itemsets that can be extracted  
(assuming minsup > 0.1)?

---

10

---

9

---

4

---

3

## Question 2

2 / 2 pts

Transaction ID Items Bought

- 1 {Milk, Beer, Diapers}
- 2 {Bread, Butter, Milk}

- 3 {Milk, Diapers, Cookies}
- 4 {Bread, Butter, Cookies}
- 5 {Beer, Cookies, Diapers}
- 6 {Milk, Diapers, Bread, Butter}
- 7 {Bread, Butter, Diapers}
- 8 {Beer, Diapers}
- 9 {Milk, Diapers, Bread, Butter}
- 10 {Beer, Cookies}

Find all itemsets (of size 3 or larger) that has the largest support.

{Milk, Beer, Diapers}

{Bread, Butter, Milk}

{Bread, Butter, Diapers}

{Milk, Diapers, Bread}

None of the above.

All of the above.

### Question 3

2 / 2 pts

Transaction ID Items Bought

- 1 {Milk, Beer, Diapers}
- 2 {Bread, Butter, Milk}
- 3 {Milk, Diapers, Cookies}
- 4 {Bread, Butter, Cookies}
- 5 {Beer, Cookies, Diapers}
- 6 {Milk, Diapers, Bread, Butter}
- 7 {Bread, Butter, Diapers}
- 8 {Beer, Diapers}
- 9 {Milk, Diapers, Bread, Butter}
- 10 {Beer, Cookies}

What is the confidence of the rule Beer  $\rightarrow$  Cookies

0.5

#### Question 4

2 / 2 pts

Apriori pruning based on support is a greedy strategy

True

False

**Question 5**

2 / 2 pts

Example of market basket transactions.

Transaction ID Items Bought

- |    |              |
|----|--------------|
| 1  | {a, b, d, e} |
| 2  | {b, c, d}    |
| 3  | {a, b, d, e} |
| 4  | {a, c, d, e} |
| 5  | {b, c, d, e} |
| 6  | {b, d, e}    |
| 7  | {c, d}       |
| 8  | {a, b, c}    |
| 9  | {a, d, e}    |
| 10 | {b, d}       |

If the minimum support is set at 40%, how many frequent 3-itemsets will be found?

## Question 6

2 / 2 pts

Select the true statements about neural network (NN) complexity.

- 
- CNN is always more complex than DNN
- 
- Adding more hidden layers increases complexity of a NN
- 

- An Autoencoder is always more complex than a CNN
- 



Complexity metric defined for a NN may not make sense for other classification engines

## Question 7

2 / 2 pts

Usage of deep learning is predicated with availability of large amounts of data. Is the data volume the only determining factor to check whether we can use deep learning ?

- 
- Yes, lower volumes lead to overfitting
- 



No, balance among classes in the input data is the only determining factor  
volume does not matter

---



No, balance among classes in the input data is a co-factor along with the  
data volume

---

- None of the above

**Question 8**

2 / 2 pts

The number of neurons in the output layer should match the number of classes (Where the number of classes is greater than 2) in a supervised learning task.

---

True

---

False

**Question 9**

2 / 2 pts

Convolutional networks played an important role in the history of deep learning as indicated by

---

they were the first deep learning models to perform well.

---

they perform better in any type of application.

---

they were used to win many machine classification contests.

---

all of the above

**Question 10**

2 / 2 pts

What is an autoencoder?

---

a neural network that copies its input to its output.

---

a neural network that codes itself.

a neural network that maps an output to an input through a hidden layer.

a neural network that is trained to attempt to copy its input to its output.

### Question 11

2 / 2 pts

Select qualities of clusters produced by a good clustering algorithm.

Intra-cluster distances are minimized

Inter-cluster distances are maximized

Number of clusters produced

A and B are correct

good cluster have strong cohesion within the cluster and maximum distance between them.

### Question 12

2 / 2 pts

Identify the number of clusters using the center-based, contiguity-based, and density-based definitions for below figure.

center-based: 1 cluster, contiguity-based: 2 clusters, density-based: 2 cluster

center-based: 2 cluster, contiguity-based: 2 clusters, density-based: 1 cluster

center-based: 3 cluster, contiguity-based: 1 clusters, density-based: 3 cluster

center-based: 2 cluster, contiguity-based: 2 clusters, density-based: 2 cluster

None of above

### Question 13

2 / 2 pts

Select different aspects of cluster validation.

Determining the clustering tendency of a set of data

To improve the complexity of the algorithm

Determining the 'correct' number of clusters

None of the above

correct these are some of the aspects of cluster validation!

### Question 14

2 / 2 pts

Select all that are not true about DBSCAN.

it has trouble when the clusters have widely varying densities

it is very sensitive to noise and to the size of the data

DBSCAN has a complexity of order  $O(n^3)$

it can be expensive when the computation of nearest neighbors requires computing all pairwise proximities

A and B are correct

B and C are correct

DBSCAN is relatively resistant to noise and can handle clusters of arbitrary shapes and sizes. It has complexity order  $O(n \log n)$

A, C and D are correct

## Question 15

2 / 2 pts

What type of clustering is DBSCAN?

- Center-based
- Contiguous
- Well-separated
- None of above

DBSCAN is a density-based clustering algorithm

Quiz Score: **30** out of 30