

# Final Exam

**Due May 8 at 4pm      Points 50      Questions 42**  
**Available May 8 at 1pm - May 8 at 4pm 3 hours      Time Limit 75 Minutes**

## Instructions

- The exam on modules 7, 8, 9, 10, 11 and 12.
- The exam will be available on Monday May 08, 2023 from 1:00 PM to 4:00 PM.
- You need to answer 38 MCQs with **1 point** for each + 4 Short questions with **3 points** for each.
- You will have only 75 minutes to complete your exam in one sitting.

## Attempt History

	Attempt	Time	Score
LATEST	<u>Attempt 1</u>	75 minutes	35 out of 50 *

\* Some questions not yet graded

❗ Correct answers will be available on May 9 at 4:30pm.

Score for this quiz: **35** out of 50 \*

Submitted May 8 at 3:40pm

This attempt took 75 minutes.

Question 1	1 / 1 pts
Why are linearly separable problems of interest of neural network researchers?	
<input type="radio"/> because they are the only class of problem that network can solve successfully	

- because they are the only mathematical functions you can draw
- because they are the only mathematical functions that are continuous
- because they are the only class of problem that perceptron can solve successfully

**Question 2**

1 / 1 pts

What is true regarding backpropagation rule?

- all of the mentioned
- 
- error in output is propagated backwards only to determine weight updates
- there is no feedback of signal at any stage
- it is also called generalized delta rule

**Question 3**

1 / 1 pts

What is perceptron?

- a single layer feed-forward neural network with pre-processing
- an auto-associative neural network
- a neural network that contains feedback
- a double layer auto-associative neural network

**Question 4****1 / 1 pts**

What is the objective of backpropagation algorithm?



to develop learning algorithm for single layer feedforward neural network



to develop learning algorithm for multilayer feedforward neural network, so that network can be trained to capture the mapping implicitly



none of the mentioned



to develop learning algorithm for multilayer feedforward neural network

**Question 5****1 / 1 pts**

How can learning process be stopped in backpropagation rule?



no heuristic criteria exist



on basis of average gradient value



there is convergence involved



none of the mentioned

**Question 6****1 / 1 pts**

Layers between the input and output layers are known as:

- Resultant layer
- Multilayer
- Hidden layer
- Output layer

**Question 7**

1 / 1 pts

Which of the following is not a correct statement for SVM.

- SVMs are resilient to overfitting
- Instances closest to the maximum margin hyperplane are called support vectors
- SVMs can be reused as algorithms for learning linear classifiers
- All instances are required to define the maximum margin hyperplane.

**Question 8**

1 / 1 pts

In SVM, the Hyper plane,  $f(x)=\text{sign}(w^*x+b)$ , where 'w' is a?

- None of the mentioned
- Distance

Vector Constant**Question 9**

1 / 1 pts

ROC in performance metrics stands for?

 Revise operating characteristic Remote operating characteristic Reverse operating characteristic Receiver operating characteristic**Question 10**

1 / 1 pts

Which of the following are real world applications of the SVM?

 All of the mentioned Image Classification Text and Hypertext Categorization Clustering of News Articles**Question 11**

1 / 1 pts

ROC chart is a \_\_\_\_\_ plot.

One-dimensional

Three-dimensional

Two-dimensional

Multi-dimensional

### Question 12

1 / 1 pts

If  $x_1, x_2$  are independent variables and  $y$  the dependent variable, which of the following represents a linear regression model?

$y = a_0 + a_1x_1 + a_2x_2$

$y = a_0 + a_1/x_1 + a_2/x_2$

$y = a_0 + a_1x_1 + a_2x_{22}$

$y = a_0 + a_1x_{12} + a_2x_2$

### Question 13

1 / 1 pts

In association rules, what is meant by the term support?

An attribute-value pair.

- Combinations of attribute-value pairs that have a minimum coverage.
- The number of instances correctly covered by the association rule.
- A frequent item set.

**Question 14**

1 / 1 pts

Which of the following best describes lift in knowledge discovery?

- A measure of interestingness of a rule
- An unsupervised learning approach
- A known class attribute
- A data mining technique

**Question 15**

1 / 1 pts

Which of the following refers to the sequence of pattern that occurs frequently?

- Frequent sub-items
- Frequent sub-structure
- All of the mentioned
- Frequent sub-sequence

**Question 16**

1 / 1 pts

Which of the following describes a strategy of frequent Itemset generation?

- Use efficient data structures to store the candidates or transactions
- Use pruning techniques to reduce the number of candidates
- 
- Reduce size of the number of transactions as the size of itemset increases
- All of the mentioned

**Question 17**

1 / 1 pts

Which of the following is not an example of frequent pattern analysis?

- Can we predict the winner of match?
- Can we automatically classify web documents?
- What kinds of DNA are sensitive to this new drug?
- What are the subsequent purchases after buying a PC?

**Question 18**

1 / 1 pts

Which of the following learning algorithm can be used to predict a combination of attributes?

- 
- Apriori.
  - Naïve Bayesian.
  - K-means
  - Decision tree.
- 

**Question 19**

1 / 1 pts

\_\_\_\_\_ is a clustering procedure characterized by the development of a tree-like structure.

- 
- K-Medoids clustering
  - Hierarchical clustering
  - K-Means clustering
  - Non-hierarchical clustering
- 

**Question 20**

1 / 1 pts

Which clustering method develops a subdivision of the given dataset into a predetermined number K of non empty subset?

- Grid method
- Hierarchical method
- Density-based method
- Partition method

**Question 21**

1 / 1 pts

Which clustering algorithm starts with each cluster comprising exactly one observation and then progressively combines the two nearest clusters until there is just one cluster left at the end?

- Agglomerative clustering
- Divisive clustering
- Single Linkage clustering
- Complete Linkage clustering

**Question 22**

1 / 1 pts

Which of the following statement is NOT true about clustering?

- It is an unsupervised learning technique
- It uses clusters for data analysis

- It is a supervised learning technique
- It groups the data

**Question 23**

1 / 1 pts

Which statement is not true about cluster analysis?

- Objects in each cluster tend to be similar to each other and dissimilar to objects in the other clusters.
- Cluster analysis is a technique for analyzing data when the criterion or dependent variable is categorical and the independent variables are interval in nature.
- Cluster analysis is also called classification analysis or numerical taxonomy.
- Groups or clusters are suggested by the data, not defined a priori.

**Question 24**

1 / 1 pts

\_\_\_\_\_ Method derives clusters from the number of observations locally falling in a neighborhood of each observation.

- Grid method

Density-based method Hierarchical method Partition method

Incorrect

**Question 25**

0 / 1 pts

Which of the following statements is incorrect about the hierachal clustering?

The choice of an appropriate metric can influence the shape of the cluster

In general, the splits and merges both are determined in a greedy manner

The hierachal type of clustering is also known as the HCA

All of the mentioned

**Question 26**

1 / 1 pts

Which of the following is not an application of cluster analysis?

Observing earth quake epicenters should be clustered along continent faults.

Identifying groups of motor insurance policy holders with a high average claim cost.

- Help marketers discover distinct groups in their customer bases.
- Decide about the subsequent purchases after buying a PC.

**Question 27**

1 / 1 pts

One of the strengths of using statistical methods in anomaly detection is:

- It can use many dimensionality reduction approaches.
- 
- Theoretically it can be applicable to all kinds of data
- It is very effective to find the outliers.
- All of the mentioned

Incorrect

**Question 28**

0 / 1 pts

One of the drawbacks of using clustering in anomaly detection is:

- Density may become less meaningful in high-dimensional space.
- Sometime it can be difficult to decide on number of clusters
- 
- It may be hard to estimate the true distribution for high dimensional data
- All of the mentioned

**Question 29**

1 / 1 pts

Which of the following will be Euclidean Distance between the two data point A(1, 3) and B(2, 3)?

---

 1

---

 4

---

 2

---

 8**Question 30**

1 / 1 pts

One of the drawbacks of using statistical methods in anomaly detection is:

---

 all of the mentioned

---

It may be hard to estimate the true distribution for high dimensional data.

---

 Sometime it can be difficult to decide on number of clusters

---

 Density may become less meaningful in high-dimensional space.**Question 31**

1 / 1 pts

Data may contain erroneous or anomalous values, which are usually referred to as \_\_\_\_\_.

- Noise
- Outliers
- Reduction
- Inconsistencies

**Question 32**

1 / 1 pts

One reason of anomaly detection is:

- All of the mentioned
- Data coming from different classes
- Normal variations can be seen on data
- Errors from collecting data

**Question 33**

1 / 1 pts

Which of the following is not type of data mining in recommender systems?

- 
- Extract from known to unknown ratings to predict user-item combinations

- All of the mentioned
- Model-based method uses a collection of ratings to learn a model
- Mining of spatiotemporal, biological, diverse semantics and relationships

**Question 34**

1 / 1 pts

Which of the following is one of the purposes of the visualization?

- It can provide qualitative overview of large data sets
- It helps find interesting regions for any further analysis.
- All of the mentioned
- It assists to search for trends and relationships among data.

**Question 35**

1 / 1 pts

Which of the following is method of preserving privacy in data mining process?

- All of the mentioned
- Add noise to the data in order to mask some attribute values of records
- Removing sensitive features or fields associated with the data
- Personal information is encrypted and stored at different locations

**Question 36**

1 / 1 pts

Which of the following is a new trend in data mining?

- Scalable data mining methods
- All of the mentioned
- Invisible data mining
- A Web mining

**Question 37**

1 / 1 pts

In web mining, \_\_\_\_\_ is used to know the order in which URLs tend to be accessed.

- sequential analysis
- associations
- classification
- clustering

Incorrect

**Question 38**

0 / 1 pts

Which of the following describes an example of the factor analysis?



For given experimental data, one analyzes the data for two or more populations described by a numeric response variable and one or more categorical variables (factors)



For certain data, researcher can indirectly measure other quantities that reflect the factor of interest



All of the mentioned



For special type of data, one attempts to determine several discriminant functions (factors) that discriminate among the groups defined by the response variable

### Question 39

Not yet graded / 3 pts

In your opinion, what are the major **5 trends** in data mining research today? Name one **major issue** in data mining, which in your view, may have a strong impact on society.

Your Answer:

Trends in data mining inquire about nowadays:

1. Deep Learning and Neural Systems: Profound learning procedures, such as convolutional neural systems (CNNs) and repetitive neural systems (RNNs), proceed to development information mining inquire about, empowering more precise and robotized design acknowledgment in expansive datasets.

2. Unsupervised Learning and Irregularity Discovery: Unsupervised learning methods, like clustering and inconsistency location, are picking up consideration as they permit the distinguishing proof of designs or irregularities in information without the require for labeled cases. This drift bolsters applications in extortion discovery, cybersecurity, and exception distinguishing proof.

3. Privacy-Preserving Information Mining: As concerns around information

security develop, analysts are centering on creating procedures that empower information mining whereas protecting person protection.

Privacy-preserving information mining strategies, counting differential protection and secure multi-party computation, point to strike a adjust between information utility and protection security.

4. Streaming Information Mining: With the expanding volume and speed of information created in real-time, there's a developing require for procedures that can productively mine and analyze spilling information. Analysts are creating calculations and frameworks that can handle nonstop information streams and extricate significant experiences in real-time.

5. Explainable AI and Interpretability: As AI models become more complex, there's a developing request for interpretability and explainability. Analysts are investigating strategies to form information mining models more straightforward and reasonable, empowering people to comprehend the thinking behind AI-driven choices.

Major issue with solid societal affect:

One major issue in information mining that might have a solid affect on society is moral contemplations and predisposition. As information mining strategies are applied in different spaces, counting back, healthcare, criminal equity, and hiring, there's a hazard of propagating inclinations or separation show within the fundamental information. One-sided information can lead to one-sided expectations or choices, possibly fortifying existing social imbalances or marginalizing certain bunches. It is pivotal to address issues of fairness, transparency, and accountability in information mining to guarantee that the benefits of these advances are conveyed evenhandedly which they don't intensify societal inclinations.

### Question 40

Not yet graded / 3 pts

For each of the following questions, provide an example of **an association rule from the market basket domain** that satisfies the following conditions. Also, describe whether such rules are subjectively interesting or not.

a) A rule that has reasonably high support but low confidence.

b) A rule that has low support and high confidence.

Your Answer:

Support means probability that antecedent is true.

Confidence means conditional probability that if antecedent is true, then with how much probability, consequent is true.

A) A rule that has reasonably high support but low confidence

MILK --> TUNA

Milk has high support. Because many people buy milk. But this rule has low confidence because , it is not necessary thing that who buys milk will also buy tuna

This type of patterns also not interested. Because it has low confidence

B) A rule that has low support and high confidence

VOTKA -->CAVIAR

This rule is very interesting because it has low support and high confidence. Votka does not appear so many times in transactions but if it appears then most of the times caviar also will appear.

this type of patterns are interesting and useful.

#### Question 41

Not yet graded / 3 pts

A database has 4 transactions, shown below.

TID	Items bought
T100	{K, A, D, B}
T200	{D, A, C, E, B}
T300	{C, A, B, E}
T400	{B, A, D}

Assuming a minimum level of support **min\_sup = 60%** and a minimum level of confidence **min\_conf = 80%**.

Given the **frequent itemsets** are:  $\{\{A\}, \{B\}, \{D\}, \{A, B\}, \{B, D\}, \{A, B, D\}\}$ .

List all of the **strong association rules** (with support  $s$  and confidence  $c$ ) **matching the following metarule**, where  $X$  is a variable representing customers, and  $item_i$  denotes variables representing items (e.g., “ $A$ ”, “ $B$ ”, etc.):

$$\forall X \in \text{transaction}, \text{buys}(X, \text{item}1) \wedge \text{buys}(X, \text{item}2) \Rightarrow \text{buys}(X, \text{item}3)[s, c]$$

Your Answer:

$\text{buys}(X, A) \wedge \text{buys}(X, B) \rightarrow \text{buys}(X, D)$  (75%, 75%) Not Strong  
 $\text{buys}(X, A) \wedge \text{buys}(X, D) \rightarrow \text{buys}(X, B)$  (75%, 100%) Strong  
 $\text{buys}(X, B) \wedge \text{buys}(X, D) \rightarrow \text{buys}(X, A)$  (75%, 100%) Strong

### Question 42

Not yet graded / 3 pts

Consider a transaction dataset that contains five items, { $A, B, C, D, E$ }. Suppose **the rules  $\{A, B\} \rightarrow C$  has the same confidence as  $\{A, B\} \rightarrow D$** , which one of the following statements are true or not, and why:

1. a) The confidence of the  $\{A, B\} \rightarrow \{C, D\}$  is the same as the confidence of  $\{A, B\} \rightarrow \{C\}$ .
2. b) All transactions that contain  $\{A, B, C\}$  also contain  $\{A, B, D\}$ .

Your Answer:

- a) The certainty of the  $\{A, B\} \rightarrow \{C, D\}$  is the same as the certainty of  $\{A, B\} \rightarrow \{C\}$ .

This explanation isn't essentially genuine. The certainty of a rule may be a degree of how frequently the ensuing (the item(s) on the right-hand side)

happens in exchanges that contain the forerunner (the item(s) on the left-hand side).

In this case, the given rules  $\{A, B\} \rightarrow C$  and  $\{A, B\} \rightarrow D$  have the same certainty, which suggests that in exchanges where A and B happen together, the event of C and D is similarly likely.

In any case, the certainty of  $\{A, B\} \rightarrow \{C, D\}$  would depend on the particular dataset and the recurrence of event of C and D together in exchanges containing A and B. In case C and D as often as possible co-occur in such exchanges, at that point the certainty of  $\{A, B\} \rightarrow \{C, D\}$  would be higher than the certainty of  $\{A, B\} \rightarrow \{C\}$ . On the other hand, in case C and D once in a while co-occur in exchanges containing A and B, at that point the certainty of  $\{A, B\} \rightarrow \{C, D\}$  would be lower than the certainty of  $\{A, B\} \rightarrow \{C\}$ .

Subsequently, the explanation isn't all around genuine and depends on the particular dataset and the conveyance of thing events.

b) All exchanges that contain  $\{A, B, C\}$  moreover contain  $\{A, B, D\}$ .

This explanation isn't fundamentally genuine. The given rules  $\{A, B\} \rightarrow C$  and  $\{A, B\} \rightarrow D$  show that when A and B happen together, C and D are conceivable consequents. In any case, it does not ensure that each exchange containing A, B, and C would too contain D. The certainty of a run the show as it were gives data approximately the probability of the ensuing given the predecessor but does not infer that all exchanges containing the forerunner will have the consequent. There may well be exchanges with A, B, and C where D does not happen.

In this manner, the articulation isn't all around genuine and does not hold based exclusively on the given rules.

Quiz Score: **35** out of 50