



STUDENT

**0002-JPD**

TENTAMEN

**TIG117 Tentamen**

---

Kurskod	--
Bedömningsform	DO
Starttid	11.05.2023 08:30
Sluttid	11.05.2023 11:30
Bedömningsfrist	--
PDF skapad	28.12.2023 14:11
Skapad av	Lisa Johansson

---

## i Information

OMTENTAMEN

**Artificiell intelligens 2: Maskininlärning och dataanalys, 7,5hp**

DAG: 11/5 -2023 TID: 08.30 – 11.30

Ansvarig: Faramarz Agahi

Förfrågningar: 031-786 28 22

Resultat: Anslås senast den 30/5 -2023

Tillåtet hjälpmedel: Tryckt svensk/engelsk ordbok (utan anteckningar)

- Totalt antal tillgängliga poäng = 30p
- Alla frågor är flervalsfrågor.
- För att få G = 16p. För att få VG = 22p.
- Antalet frågor 30 (1 p per fråga)
- Frågor kräver antingen ett svar eller flera svar.
- Om flera svar krävs är det nödvändigt att få svaret helt korrekt för att få poängen (det är allt eller inget!).

### 1 Q1

Vilket påstående är sant?

**Välj ett alternativ:**

- Enligt David Hume går all kunskap tillbaka på logik
- Ada Lovelace var först med att programmera en maskin



Rätt. 1 av 1 poäng.

**2 Q2**

Vem skrev "Nothing in Biology Makes Sense Except in the Light of Evolution"?

**Välj ett alternativ:**

- John McCarthy
- Theodosius Dobzhansky
- Charles Darwin



Rätt. 1 av 1 poäng.

**3 Q3**

Vilket påstående är sant?

**Välj ett alternativ:**

- Man känner inte till några effektiva algoritmer för handelsresandeproblemet
- Man har numera tillgång till effektiva algoritmer för handelsresandeproblemet



Rätt. 1 av 1 poäng.

**4 Q4**

Vilket påstående är sant?

**Välj ett alternativ:**

- En violin plot innehåller mindre information än en box plot
- En quiver plot använder pilar för att visualisera flöden



Rätt. 1 av 1 poäng.

**5 Q5**

Vilket påstående är sant?

**Välj ett alternativ:**

- Tiodimensionella dataset kan representeras som heat maps med tio färger
- Tiodimensionella dataset kan representeras i ett tvådimensionellt kalkylark 

Rätt. 1 av 1 poäng.

**6 Q6**

Vilket påstående är sant?

**Välj ett alternativ:**

- Navigatorer bygger på sökalgoritmer 
- Navigatorer bygger på maskininlärning

Rätt. 1 av 1 poäng.

**7 Q7**

Vilket påstående är sant?

**Välj ett alternativ:**

- Som regel är testmängden större än träningsmängden
- Bootstrapping används för att skapa variation 
- Beslutsträd kan inte hantera numeriska features

Rätt. 1 av 1 poäng.

**8 Q8**

Vilket påstående är sant?

**Välj ett alternativ:**

- Overfitting innebär att modellen är för grov
- Med en confusion matrix kan man studera overfitting
- Vid stora värden på  $k$  tenderar  $k$ -nn att ge overfitting



Rätt. 1 av 1 poäng.

**9 Q9**

Vilket påstående är sant?

**Välj ett alternativ:**

- L1 normen mäter det euklidiska avståndet
- L1 normen kallas även Manhattanavståndet



Rätt. 1 av 1 poäng.

**10 Q10**

Vilket påstående är sant?

**Välj ett alternativ:**

- Beslutsträd är ett exempel på en ensemble-metod
- En random forest är enklare att tolka än ett beslutsträd
- Random forest bygger på att man väljer det bästa beslutsträdet



Fel. 0 av 1 poäng.

**11 Q11**

Vilket påstående är sant?

**Välj ett alternativ:**

- Random forest har i allmänhet större problem med overfitting än beslutsträd
- Random forest använder slump till både bootstrapping och restriktioner på träden 
- Bootstrapping av ett dataset D kan leda till mängder som innehåller fler element än D

Rätt. 1 av 1 poäng.

**12 Q12**

Vilket påstående är sant?

**Välj ett alternativ:**

- Regularisering används för att minska overfitting 
- Outliers sorteras automatiskt bort av regularisering
- Regularisering innebär att man försöker reducera prediktionsfelen

Rätt. 1 av 1 poäng.

**13 Q13**

Vilket påstående är sant?

**Välj ett alternativ:**

- I DBSCAN anger man på förhand hur många kluster man vill ha
- Klustering är en form av reinforcement learning
- Klustering är en form av unsupervised learning



Rätt. 1 av 1 poäng.

**14 Q14**

Vilket av dessa påståenden om k-means klustering är korrekt?

**Välj ett alternativ:**

- det är en metod för att konstruera k kluster
- den bygger på att man successivt flyttar datapunkter närmare centralpunkterna
- det är en metod för klustering i maximalt k dimensioner



Rätt. 1 av 1 poäng.

**15 Q15**

Vilket påstående är sant?

**Välj ett alternativ:**

- Alla punkter kommer med i något kluster när man använder DBSCAN
- Hierarkisk klustring tar inte med datapunkter som är outliers
- Resultatet av hierarkisk klustring är ett dendrogram



Rätt. 1 av 1 poäng.

**16 Q16**

Supervised learning and unsupervised clustering both require which is correct according to the statement.

**Välj ett alternativ:**

- hidden attribute
- input attribute
- output attribute



Rätt. 1 av 1 poäng.

## 17 Q17

The output of training process in machine learning is \_\_\_\_\_

Välj ett alternativ:

- machine learning model
- null
- accuracy



Rätt. 1 av 1 poäng.

## 18 Q18

Below is a truth table for a simple classification problem.

Object	Purple?	Round?	Eat?
Blueberry	1	1	1
Golf ball	0	1	0
Violet	1	0	0
Hot Dog	0	0	1

Four labelled objects are required to be classified as eatable/edible or not. The ANN receives as inputs a two-element vector of the object properties "purple", "round" and must make the decision ("eat" classification) based on those properties. Which of the following are true:

**More than one correct answer:**

- The problem requires one (or more) hidden layers to solve.
- The input-output mapping is linearly separable.
- The problem can be viewed as a (Not) XOR problem (inverse of XOR problem).



Rätt. 1 av 1 poäng.

**19 Q19**

How many weights (excluding threshold/bias weights) does a fully-connected/dense ANN with two hidden layers consisting of 3 input nodes, 2 nodes in hidden layer 1, 2 nodes in hidden layer 2, and 1 output node, have?

**Only one of the below is correct:**

- 8
- 10
- 12
- 18



Rätt. 1 av 1 poäng.

**20 Q20**

Given an input vector of  $\mathbf{x} = [x_1, x_2]$  where  $x_1 = 0.25$ ,  $x_2 = 1.0$ , and a weight vector  $\mathbf{w} = [1.0, 0.5]$ , what is the activation value as calculated by the dot product of  $\mathbf{x}$  and  $\mathbf{w}$ .

**One correct answer:**

- [0.25, 0.5]
- 2.75
- 0.75
- [1.25, 1.5]



Rätt. 1 av 1 poäng.

**21 Q21**

How many pooling maps does a standard 'shallow' Convolutional Neural Network (1 layer of feature maps + 1 layer of pooling maps) have that uses 4 filters of dimensionality 3x3 over the input image?

**Choose one from the below**

4



9

54

14

Rätt. 1 av 1 poäng.

**22 Q22**

Which of the below correctly describe similarities and differences between convolution filters (used for encoding on feature maps) and pooling "filters" (used for encoding on pooling maps)?

**Choose one from below**

Convolution filters are used for dot product computations and pooling filters are used for max or mean computations



Convolution filters are used for max or mean computations and pooling filters are used for dot product computations

Only convolutional filters 'scan' over their processed inputs

Both types of filter have trainable weights

Rätt. 1 av 1 poäng.

**23 Q23**

Of the following, which is the **best** characterization of the Vanishing Gradient Problem

**Choose one from below**

- A problem for backpropagating error terms where the ANN is shallow
- Backpropagating error/loss leads to increasingly small values for updating the weights nearer to the input layer of the neural network. 

Rätt. 1 av 1 poäng.

**24 Q24**

Which of the below best describes the term "local minimum" as compared to "global minimum"?

**Choose one from below**

- A point at which the network loss is at its absolute minimum
- A point at which the network loss is not at its potential minimum 

Rätt. 1 av 1 poäng.

**25 Q25**

Which of the following make CNNs efficient with respect to the number of parameters they have as compared to "dense" ANNs?

**Välj ett alternativ:**

- Each feature map consists of a number of neurons whose encodings share a single filter of weights 
- Adding convolutional layers to CNN always means having fewer parameters (weights) in the full network

Rätt. 1 av 1 poäng.

**26 Q26**

Which of the following are considered good stopping criteria when training ANNs?

**Välj ett alternativ:**

- Successive non decreasing loss/error for validation data checks
- Validation loss/error below validation loss



Rätt. 1 av 1 poäng.

**27 Q27**

How can you handle missing or corrupted data in a dataset?

**Välj ett alternativ:**

- Drop missing rows or columns
- All of the above
- Replace missing values with mean/median/mode



Rätt. 1 av 1 poäng.

**28 Q28**

Machine learning algorithms build a model based on sample data, known as

.....

**Välj ett alternativ:**

- Training Data
- None of the above
- Data Training
- Transfer Data



Rätt. 1 av 1 poäng.

**29 Q29**

Common classes of problems in machine learning is .....

**Välj ett alternativ:**

- Classification
- Regression
- Clustering
- All of the above



Rätt. 1 av 1 poäng.

**30 Q30**

Choose one from the below that best characterizes supervised learning?

**Välj ett alternativ:**

- Learning that compares ANN predictions about data to human labelled data
- Learning that is based on whether or not the output of the network gives a value (or values) of 1.
- Learning that compares ANN predictions about data to non-human labelled data
- Learning that is based on similarities between input values and output values



Rätt. 1 av 1 poäng.