Technische Hochschule Rosenheim

Faculty of Computer Science
Prof. Dr. M. Tilly

Winter Term 2022/2023

Mock Exam - Applied Artificial Intelligence(AAI)
400 - Unsupervised and Reinforcement Learning(URL)

Date: 02.01.2023	Duration: 90 Minutes	Material: handwritten notes A4; calculator

---

Name: ....

Matrix number: ....

Good luck!

Notes:

1. The staples must not be loosened. The exam includes **16 pages incl. cover sheet and worksheets.**.
2. Work on the questions directly in the task. If necessary, use the worksheets at the end.
3. If, in your opinion, there are contradictions in the tasks or information is missing, make reasonable assumptions and document them..
4. The distribution of points is for orientation, but it is not binding.
5. Please do not write in pencil, red or green pens and if possible **legible**.

SOLUTION is available in the new year!

Name: Matrix number:

Worksheet

Name: Matrix number:

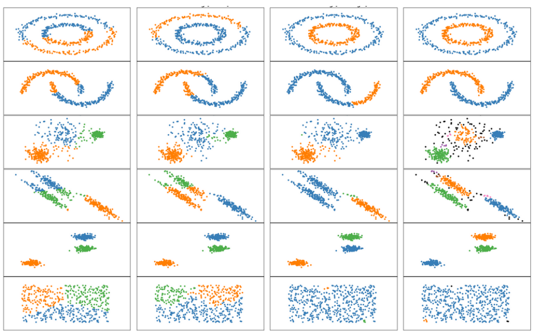
1. Task - General questions 4+6 Points

a)

Mark the correct answer or statement; mark exactly one answer per question.

1. Clustering
   - ☐ The silhouette score calculates the surface of a silhouette for eigenfaces,
   - ☐ The elbow method can be used to determine the number of clusters k in a given data set.
   - ☐ The elbow method is sometimes also called the KneeLocator.
2. k-means vs. DBSCAN
   - ☐ For k-means the number of clusters is calculated by the algorithm.
   - ☐ k-means is faster the DBSCAN.
   - ☐ DBSCAN does not need a-priori knowledge on number of clusters.
3. Clustering

   Given the following figure. Each column shows the result of a cluster algorithm for a different data set. Name 2 out of 4 clustering algorithms. Write on top of each column:



4. Pre-Processing and Normalization

   Name 2 normalization methods:

b)

Describe the DBSCAN algorithm.

2. Task - k-means Clustering 15 Points

Use the k-means algorithm and Euclidean distance to cluster the following 8 examples into 3 clusters:

A1=(2,10), A2=(2,5), A3=(8,4), A4=(5,8), A5=(7,5), A6=(6,4), A7=(1,2), A8=(4,9).

The distance matrix based on the Euclidean distance is given below

	A1	A2	A3	A4	A5	A6	A7	A8
A1	0	$\sqrt{25}$	$\sqrt{36}$	$\sqrt{13}$	$\sqrt{50}$	$\sqrt{52}$	$\sqrt{65}$	$\sqrt{5}$
A2		0	$\sqrt{37}$	$\sqrt{18}$	$\sqrt{25}$	$\sqrt{17}$	$\sqrt{10}$	$\sqrt{20}$
A3			0	$\sqrt{25}$	$\sqrt{2}$	$\sqrt{2}$	$\sqrt{53}$	$\sqrt{41}$
A4				0	$\sqrt{13}$	$\sqrt{17}$	$\sqrt{52}$	$\sqrt{2}$
A5					0	$\sqrt{2}$	$\sqrt{45}$	$\sqrt{25}$
A6						0	$\sqrt{29}$	$\sqrt{29}$
A7							0	$\sqrt{58}$
A8								0

Suppose that the initial centers of each cluster are A1, A4 and A7. **Run the k-means algorithm for 1 epoch only**. At the end of this epoch show:

1. The new clusters (i.e. the examples belonging to each cluster)
2. The centers of the new clusters
3. Draw a 10 by 10 space with all the 8 points and show the clusters after the first epoch and the new centroids

Name: Matrix number:

Worksheet Clustering

Name: Matrix number:

3. Task - PCA **2+6+2+5 Points**

Consider the following data points (2, 1), (3, 5), (4, 3), (5, 6), (6, 7), (7, 8).
Compute the principal component using PCA Algorithm by the following steps:

1. Calculate the mean vector.
2. Calculate the covariance matrix.
3. Proof that $\lambda 1 = 8.22$ and $\lambda 2 = 0.38$ are eigenvalues of the covariance matrix.
4. Calculate the PCA (or eigenvector).

Name: Matrix number:

Worksheet PCA

4. Task - DBSCAN **10+5 Points**

If Epsilon is 2 and minpoint is 2 (minpts), what are the clusters that DBScan would discover with the following 8 examples: **A1=(2,10), A2=(2,5), A3=(8,4), A4=(5,8), A5=(7,5), A6=(6,4), A7=(1,2), A8=(4,9)**.

The distance matrix is the following:

	A1	A2	A3	A4	A5	A6	A7	A8
A1	0	$\sqrt{25}$	$\sqrt{36}$	$\sqrt{13}$	$\sqrt{50}$	$\sqrt{52}$	$\sqrt{65}$	$\sqrt{5}$
A2		0	$\sqrt{37}$	$\sqrt{18}$	$\sqrt{25}$	$\sqrt{17}$	$\sqrt{10}$	$\sqrt{20}$
A3			0	$\sqrt{25}$	$\sqrt{2}$	$\sqrt{2}$	$\sqrt{53}$	$\sqrt{41}$
A4				0	$\sqrt{13}$	$\sqrt{17}$	$\sqrt{52}$	$\sqrt{2}$
A5					0	$\sqrt{2}$	$\sqrt{45}$	$\sqrt{25}$
A6						0	$\sqrt{29}$	$\sqrt{29}$
A7							0	$\sqrt{58}$
A8								0

1. Draw the 10 by 10 space and illustrate the discovered clusters.
2. What if Epsilon is increased to 10?

Name: Matrix number:

Worksheet PCA

5. Task - Autoencoder 4+6 Points

a)

Given the following code.

```
class Denoise(Model):
   def __init__(self):
      super(Denoise, self).__init__()
      self.encoder = tf.keras.Sequential([
      layers.InputLayer(input_shape=(28, 28, 1)),
      layers.Conv2D(16, (3, 3), activation='relu', padding='same', strides
         =2),
      layers.Conv2D(8, (3, 3), activation='relu', padding='same', strides
         =2)])

      self.decoder = tf.keras.Sequential([
      layers.Conv2DTranspose(8, kernel_size=3, strides=2, activation='relu'
         , padding='same'),
      layers.Conv2DTranspose(16, kernel_size=3, strides=2, activation='relu
         ', padding='same'),
      layers.Conv2D(1, kernel_size=(3, 3), activation='sigmoid', padding='
         same')])

   def call(self, x):
      encoded = self.encoder(x)
      decoded = self.decoder(encoded)
      return decoded
}
```

Provide the shape of each layer:

b)

Explain how to use **Autoencoder** for anomaly detection!

Name: Matrix number:

6. Task - Reinforcement Learning 3+5+10 Points

a)

Explain the concept to *Reinforcement Learning* using a sketch.

b)

List and describe the characteristics of a Markov Decision Process (MDP).

c)

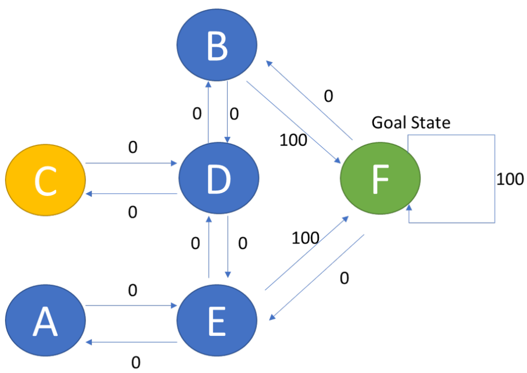
Q-learning is a variant of making an agent experience an environment without knowing the model behind it (learning without knowledge of the MDP). With Q-learning the explicit learning of the *policy* is omitted, instead the *policy* is learned directly.
The following formula applies:

$$Q_{k+1}(s_t, a_t) < -Q_k(s_t, a_t) + \alpha[R(s,a) + \gamma \max(Q_k(s_{t+1}, a)) - Q_k(s_t, a_t))]$$

where α is the *learning rate* and γ is the *discount factor*.
In this example, let the reward matrix R be given as:

state/action	A	B	C	D	E	F
A	-	-	-	-	0	-
B	-	-	-	0	-	100
C	-	-	-	0	-	-
D	-	0	0	-	0	-
E	0	-	-	0	-	100
F	-	0	-	-	0	100

As can be easily seen, there are 6 states $S = A, B, C, D, E, F$ and the actions A that allow an agent to move from a state S_1 to a state S_2 (e.g. from A to E or from D to B,C or E). The example could be a building with rooms, and doors that allow an agent to transition from one room to another.



Apply the Q-learning algorithm step by step. Calculate the following values (k denotes the respective episodes) with $\alpha = 1$ and $\gamma = 0.8$. An episode ends when the goal (=goal state) is reached:

$Q_{k=1}(B, F) =$

$Q_{k=2}(D, B) =$

$Q_{k=3}(C, D) =$

$Q_{k=4}(E, D) =$

What does the $Q - matrix$ look like after the 4 episodes?

$Q =$

Name: Matrix number:

Worksheet