IN-CLASS EXERCISE (I5)

Student ID:..............................

Duration: 15 mins   Date: 16/08/2023

Student name: ....

Score: .... / 3

Question 1 (2pts) A robot in a lumber yard learns to discriminate Oak wood from Pine wood by observing the appearance of a given piece of wood. The aside table shows the training data set.

#	Density	Grain	Hardness	Class
1	Light	Small	Hard	Oak
2	Heavy	Large	Hard	Oak
3	Light	Large	Hard	Oak
4	Heavy	Small	Medium	Oak
5	Light	Small	Medium	Pine
6	Heavy	Large	Soft	Pine
7	Light	Large	Soft	Pine
8	Heavy	Small	Soft	Pine

a) (1.5pts) Use **ID3 algorithm** to build a decision tree from the given data. Fill in the following tables with your numerical results, each of which is rounded to three decimal places. Attributes having the same best metric values are chosen in alphabetical order.

Evaluate all attributes and circle the attribute that is selected for the root node of the ID3 decision tree.

	Whole dataset	Density		Grain		Hardness		
		Heavy	Light	Large	Small	Hard	Medium	Soft
Entropy								
Average Entropy								
Information Gain								

For the branch that still contains a mixture of positive and negative examples, evaluate the remaining attributes and circle the attribute that is selected for the next node (you may leave irrelevant cells blank).

	Whole subset	Density		Grain		Hardness		
		Heavy	Light	Large	Small	Hard	Medium	Soft
Entropy								
Average Entropy								
Information Gain								

b) (0.5pt) Draw the complete decision tree.

Question 2 (1pt) Identify which type of learning should be used to solve the following task. Explain your choice. *The farmer is currently in his durian orchard, which contains two varieties of durians: "Musang King" and "6 Ri". His current task involves arranging the durians, getting them ready for sale, as each variety commands a distinct price. He determines the type of durian by observing the fruits' shapes, colors, and spikes.*

....

....

....

IN-CLASS EXERCISE (I5)

Student ID:...............................

Duration: 15 mins   Date: 16/08/2023

Student name: ....

Score: ..../ 3

Question 1 (2pts) We have some data about when people go hiking. The data takes into effect, whether the hike is on a weekend or not, if the weather is rainy or sunny, and if the person has company during the hike.

#	Weekend?	Company?	Weather	Go Hiking?
1	Y	N	R	N
2	Y	Y	R	N
3	Y	Y	S	Y
4	Y	N	S	Y
5	Y	Y	R	N
6	Y	Y	S	Y
7	N	Y	S	N
8	N	Y	R	N
9	N	N	S	N

a) (1.5pts) Use **ID3 algorithm** to build a decision tree from the given data. Fill in the following tables with your numerical results, each of which is rounded to three decimal places. Attributes having the same best metric values are chosen in alphabetical order.

Evaluate all attributes and circle the attribute that is selected for the root node of the ID3 decision tree.

	Whole	Company?		Weather		Weekend?	
	dataset	N	Y	R	S	N	Y
Entropy							
Average Entropy							
Information Gain							

For the branch that still contains a mixture of positive and negative examples, evaluate the remaining attributes and circle the attribute that is selected for the next node (you may leave irrelevant cells blank).

	Whole	Company?		Weather		Weekend?	
	subset	N	Y	R	S	N	Y
Entropy							
Average Entropy							
Information Gain							

b) (0.5pt) Draw the complete decision tree.

Question 2 (1pt) Identify which type of learning should be used to solve the following task. Explain your choice. *A marketing representative is examining the customer data, aiming to pinpoint clusters of customers who exhibit similar behaviors. The number of clusters is inherent to the data and cannot be predetermined. This analysis enhances the efficacy of customer outreach for introducing new marketing campaigns.*

....

....

....

IN-CLASS EXERCISE (I5)

Student ID: ....

Student name: ....

Duration: 15 mins

Date: 16/08/2023

Score: .... / 3

Question 1 (2pts) A student is considering whether to go to the party or not. He is thinking about whether he may have a terrible hangover the next morning, whether the party is held at weekend, and how difficult the incoming exam is.

a) (1.5pts) Use **ID3 algorithm** to build a decision tree from the given data. Fill in the following tables with your numerical results, each of which is rounded to three decimal places. Attributes having the same best metric values are chosen in alphabetical order.

#	Hangover	Exam	Weekend	Party
1	No	Easy	No	Yes
2	No	Hard	No	No
3	No	No	No	Yes
4	No	No	Yes	Yes
5	Yes	Easy	No	No
6	Yes	Hard	No	No
7	Yes	No	No	No
8	Yes	No	Yes	No

Evaluate all attributes and circle the attribute that is selected for the root node of the ID3 decision tree.

	Whole dataset	Hangover		Weekend		Exam		
		No	Yes	No	Yes	Easy	Hard	No
Entropy								
Average Entropy								

For the branch that still contains a mixture of positive and negative examples, evaluate the remaining attributes and circle the attribute that is selected for the next node (you may leave irrelevant cells blank).

	Whole subset	Hangover		Weekend		Exam		
		No	Yes	No	Yes	Easy	Hard	No
Entropy								
Average Entropy								

b) (0.5pt) Draw the complete decision tree.

Question 2 (1pt) Identify which type of learning should be used to solve the following task. Explain your choice. *You are at the final stage of a video game. The boss is so mighty that you cannot defeat it by directly facing it. After several trials, you have found that the boss has some death points. You focus on those points throughout the battle, and then you are done.*

....

....

....

IN-CLASS EXERCISE (I5)

Student ID:.................................

Duration: 15 mins   Date: 16/08/2023

Student name: ....

Score: ..../ 3

Question 1 (2pts) A robot in a fruit farm learns to determine whether an avocado is Tasty or Bad by observing the avocado's appearance. The aside table shows the training data set.

a) (1.5pts) Use **ID3 algorithm** to build a decision tree from the given data. Fill in the following tables with your numerical results, each of which is rounded to three decimal places. Attributes having the same best metric values are chosen in alphabetical order.

#	Shape	Skin	Hardness	Class
1	Round	Rough	Soft	Tasty
2	Round	Rough	Hard	Tasty
3	Oval	Smooth	Soft	Tasty
4	Oval	Rough	Hard	Tasty
5	Long	Rough	Soft	Bad
6	Round	Smooth	Soft	Bad
7	Round	Smooth	Hard	Bad
8	Long	Rough	Soft	Bad

Evaluate all attributes and circle the attribute that is selected for the root node of the ID3 decision tree.

	Whole dataset	Skin		Hardness		Shape		
		Smooth	Rough	Hard	Soft	Long	Oval	Round
Entropy								
Average Entropy								
Information Gain								

For the branch that still contains a mixture of positive and negative examples, evaluate the remaining attributes and circle the attribute that is selected for the next node (you may leave irrelevant cells blank).

	Whole subset	Skin		Hardness		Shape		
		Smooth	Rough	Hard	Soft	Long	Oval	Round
Entropy								
Average Entropy								
Information Gain								

b)   (0.5pt) Draw the complete decision tree.

Question 2 (1pt) Both supervised learning and reinforcement learning have feedback to the learner. Point out the difference between two types of feedback and thus how this difference affects the learning process.

....

....

....

....

SOLUTION

IN-CLASS EXERCISE (I5)

Student ID:.......................................

Duration: 15 mins

Date: 12/04/2023

Student name: ....

Score: ..../ 3

Question 1 (2pts) A robot in a lumber yard learns to discriminate Oak wood from Pine wood by observing the appearance of a given piece of wood. The aside table shows the training data set.

a) (1.5pts) Use **ID3 algorithm** to build a decision tree from the given data. Fill in the following tables with your numerical results, each of which is rounded to three decimal places. Attributes having the same best metric values are chosen in alphabetical order.

#	Density	Grain	Hardness	Class
1	Light	Small	Hard	Oak
2	Heavy	Large	Hard	Oak
3	Light	Large	Hard	Oak
4	Heavy	Small	Medium	Oak
5	Light	Small	Medium	Pine
6	Heavy	Large	Soft	Pine
7	Light	Large	Soft	Pine
8	Heavy	Small	Soft	Pine

Evaluate all attributes and circle the attribute that is selected for the root node of the ID3 decision tree.

	Whole dataset	Density		Grain		Hardness		
		Heavy	Light	Large	Small	Hard	Medium	Soft
Entropy	1	1	1	1	1	0	1	0
Average Entropy		1		1		0.25		
Information Gain		0		0		0.75		

For the branch that still contains a mixture of positive and negative examples, evaluate the remaining attributes and circle the attribute that is selected for the next node (you may leave irrelevant cells blank).

	Whole subset	Density		Grain		Hardness		
		Heavy	Light	Large	Small	Hard	Medium	Soft
Entropy	1	0	0	0	1			
Average Entropy		0		1				
Information Gain		1		0				

b) (0.5pt) Draw the complete decision tree.

Hardness = Hard: Class = Oak

Hardness = Soft: Class = Pine

Hardness = Medium

|-------- Density = Heavy: Class = Oak

|-------- Density = Light: Class = Pine

Question 2 (1pt) Identify which type of learning should be used to solve the following task. Explain your choice. *The farmer is currently in his durian orchard, which contains two varieties of durians: "Musang King" and "6 Ri". His current task involves arranging the durians, getting them ready for sale, as each variety commands a distinct price. He determines the type of durian by observing the fruits' shapes, colors, and spikes.*

Supervised learning. The number of classes is known in advanced, which are "Musang King" and "6 Ri". The objects in each class are distinguished by their shapes, colors, and spikes. ....

IN-CLASS EXERCISE (I5)

Student ID:................................

Duration: 15 mins   Date: 20/04/2023

Student name: ....

Score: ..../ 3

Q1 (2pts) We have some data about when people go hiking. The data takes into effect, whether the hike is on a weekend or not, if the weather is rainy or sunny, and if the person has company during the hike.

#	Weekend?	Company?	Weather	Go Hiking?
1	Y	N	R	N
2	Y	Y	R	N
3	Y	Y	S	Y
4	Y	N	S	Y
5	Y	Y	R	N
6	Y	Y	S	Y
7	N	Y	S	N
8	N	Y	R	N
9	N	N	S	N

a) (1.5pts) Use **ID3 algorithm** to build a decision tree from the given data. Fill in the following tables with your numerical results, each of which is rounded to three decimal places. Attributes having the same best metric values are chosen in alphabetical order.

Evaluate all attributes and circle the attribute that is selected for the root node of the ID3 decision tree.

	Whole dataset	Company?		Weather		Weekend?	
		N	Y	R	S	N	Y
Entropy	0.918	0.918	0.918	0	0.971	0	1
Average Entropy		0.918		0.539		0.667	
Information Gain		0		0.379		0.252	

For the branch that still contains a mixture of positive and negative examples, evaluate the remaining attributes and circle the attribute that is selected for the next node (you may leave irrelevant cells blank).

	Whole subset	Company?		Weather		Weekend?	
		N	Y	R	S	N	Y
Entropy	0.971	1	0.918			0	0
Average Entropy		0.951				0	
Information Gain		0.2				0.971	

b) (0.5pt) Draw the complete decision tree.

Weather = S
|------- Weekend? = No: No
|------- Weekend? = Yes: Yes
Weather = R: No

Question 2 (1pt) Identify which type of learning should be used to solve the following task. Explain your choice. *A marketing representative is examining the customer data, aiming to pinpoint clusters of customers who exhibit similar behaviors. The number of clusters is inherent to the data and cannot be predetermined. This analysis enhances the efficacy of customer outreach for introducing new marketing campaigns.*

Unsupervised learning. The number of clusters cannot be predetermined. Customers of the same cluster have similar behaviors, while those of different behaviors are separated to other clusters. ....

IN-CLASS EXERCISE (I5)

Student ID:...................................

Student name: ....

Duration: 15 mins

Date: 12/04/2023

Score: ..../ 3

Question 1 (2pts) A robot in a fruit farm learns to determine whether an avocado is Tasty or Bad by observing the avocado's appearance. The aside table shows the training data set.

a) (1.5pts) Use **ID3 algorithm** to build a decision tree from the given data. Fill in the following tables with your numerical results, each of which is rounded to three decimal places. Attributes having the same best metric values are chosen in alphabetical order.

#	Shape	Skin	Hardness	Class
1	Round	Rough	Soft	Tasty
2	Round	Rough	Hard	Tasty
3	Oval	Smooth	Soft	Tasty
4	Oval	Rough	Hard	Tasty
5	Long	Rough	Soft	Bad
6	Round	Smooth	Soft	Bad
7	Round	Smooth	Hard	Bad
8	Long	Rough	Soft	Bad

Evaluate all attributes and circle the attribute that is selected for the root node of the ID3 decision tree.

	Whole dataset	Skin		Hardness		Shape		
		Smooth	Rough	Hard	Soft	Long	Oval	Round
Entropy	1	0.918	0.971	0.918	0.971	0	0	1
Average Entropy		0.951		0.951		0.5		
Information Gain		0.049		0.049		0.5		

For the branch that still contains a mixture of positive and negative examples, evaluate the remaining attributes and circle the attribute that is selected for the next node (you may leave irrelevant cells blank).

	Whole subset	Skin		Hardness		Shape		
		Smooth	Rough	Hard	Soft	Long	Oval	Round
Entropy	1	0	0	1	1			
Average Entropy		0		1				
Information Gain		1		0				

b) (0.5pt) Draw the complete decision tree.

Shape = Oval: Class = Tasty

Shape = Long: Class = Bad

Shape = Round

|-------- Skin = Rough: Class = Tasty

|-------- Skin = Smooth: Class = Bad

Question 2 (1pt) Identify which type of learning should be used to solve the following task. Explain your choice. *You are at the final stage of a video game. The boss is so mighty that you cannot defeat it by directly facing it. After several trials, you have found that the boss has some death points. You focus on those points throughout the battle, and then you are done.*

Reinforcement learning. It is a trial-and-error process in which the agent only receives signal indicating success or failure and itself determines how to update the policy (no specific measurement of how to adjust the error)..

IN-CLASS EXERCISE (I5)

Student ID:.....................................

Duration: 15 mins   Date: 20/04/2023

Student name: ....

Score: .... / 3

Q1 (2pts) A student is considering whether to go to the party or not. He is thinking about whether he may have a terrible hangover the next morning, whether the party is held at weekend, and how difficult the incoming exam is.

#	Hangover	Exam	Weekend	Party
1	No	Easy	No	Yes
2	No	Hard	No	No
3	No	No	No	Yes
4	No	No	Yes	Yes
5	Yes	Easy	No	No
6	Yes	Hard	No	No
7	Yes	No	No	No
8	Yes	No	Yes	No

a) (1.5pts) Use **ID3 algorithm** to build a decision tree from the given data. Fill in the following tables with your numerical results, each of which is rounded to three decimal places. Attributes having the same best metric values are chosen in alphabetical order.

Evaluate all attributes and circle the attribute that is selected for the root node of the ID3 decision tree.

	Whole dataset	Hangover		Weekend		Exam		
		No	Yes	No	Yes	Easy	Hard	No
Entropy	0.954	0.811	0	0.918	1	1	0	1
Average Entropy		0.406		0.939		0.75		
		0.548		0.015		0.204		

For the branch that still contains a mixture of positive and negative examples, evaluate the remaining attributes and circle the attribute that is selected for the next node (you may leave irrelevant cells blank).

	Whole subset	Hangover		Weekend		Exam		
		No	Yes	No	Yes	Easy	Hard	No
Entropy	0.811			0.918	0	0	0	0
Average Entropy				0.689		0		
				0.123		0.811		

b) (0.5pt) Draw the complete decision tree.

Hangover = No
|------- Exam = Easy: Yes
|------- Exam = Hard: No
|------- Exam = No: No
Hangover = Yes: No

Question 2 (1pt) Both supervised learning and reinforcement learning have feedback to the learner. Point out the difference between two types of feedback and thus how this difference affects the learning process.

Difference: SL can precisely adjust the model based on the difference between the actual value produced by the model and the ground truth, while the feedback from the environment in RL is quite simple and not informative enough to calculate the adjustments ....