IN-CLASS EXERCISE (I5)

Student ID:..... Duration: 15 mins Date: 18/04/2023

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.

#	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.

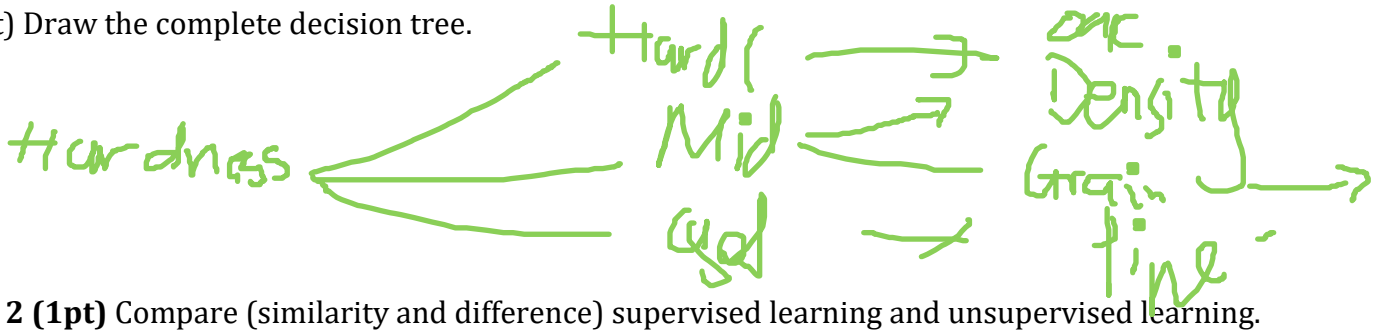
Evaluate all attributes and circle the attribute that is selected for the root node of the ID3 decision tree.

	Whole dataset	Density		Grain		Hardness		
		Heavy	Light	Large	Small	Hard	Medium	Soft
Entropy	1	1	1	1	1	1	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.



Question 2 (1pt) Compare (similarity and difference) supervised learning and unsupervised learning.

.....

.....

.....

.....

IN-CLASS EXERCISE (I5)

Student ID:..... Duration: 15 mins Date: 18/04/2023

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.

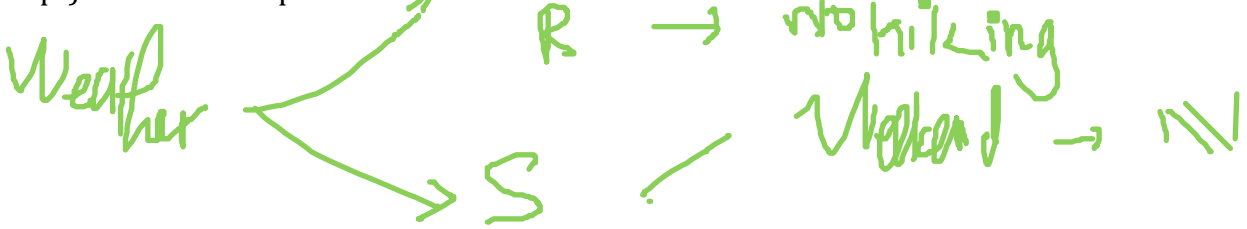
Evaluate all attributes and circle the attribute that is selected for the root node of the ID3 decision tree.

	Whole dataset	Company?		Weather		Weekend?	
		N	Y	R	S	N	Y
Entropy	0,918	0	0,918	0	0,971	0	1
Average Entropy		0,918		0,539		0,667	
Information Gain		0		0,379		0,251	

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		0,02				0,971	

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



Question 2 (1pt) Compare (similarity and difference) supervised learning and reinforcement learning.

.....

.....

.....

.....

IN-CLASS EXERCISE (I5)

Student ID:..... Duration: 15 mins Date: 18/04/2023

Student name: ..... 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.

#	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								
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) Give an example of the reinforcement learning problem. Describe the problem context. State clearly the rewards and how the agent uses that information to improve its performance.

.....

.....

.....

IN-CLASS EXERCISE (I5)

Student ID:..... Duration: 15 mins Date: 18/04/2023

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.

#	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

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	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) What is a hypothesis? Why is it that we choose the simplest consistent hypothesis?

.....

.....

.....

.....

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.

#	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	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) Compare (similarity and difference) supervised learning and unsupervised learning.

Similarity: both requires training data to build the mode.....

Difference: SL requires labeled training data while USL works on unlabeled data. SL generalizes the mapping from features to labels, while USL groups data points based on their features.....

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

Q2 (1pt) Compare (similarity and difference) supervised learning and reinforcement learning.

Similarity: both has the feedback information, SL is guided by the difference between the actual value produced by the model and the ground truth, while RL is guided by the feedback from the environment

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 in RL is quite simple and not informative enough

IN-CLASS EXERCISE (I5)

Student ID:..... Duration: 15 mins Date: 12/04/2023

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.

#	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

c) (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	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				

d) (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) Give an example of the reinforcement learning problem. Describe the problem context. State clearly the rewards and how the agent uses that information to improve its performance.

You are training your dog to get the stick. Each time the dog returns a stick successfully, you offer it favorite treats. Eventually, the dog understands that whenever the master throws a stick, it should get it as early as possible to gain a reward (a bone) from a master in a lesser time.

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

c) (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		

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

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

Q2 (1pt) What is a hypothesis? Why is it that we choose the simplest consistent hypothesis?

A hypothesis gives an estimation of the true underlying distribution of data, and it is needed to be verified. The simplest consistent hypothesis can fit all the training data points while it is general enough to avoid overfitting