

I pledge my honor that I have abided by the Stevens Honor System.

Your signature:

Date:

1. Answer the following concept questions (please make your answers concise):

1) [5 pts] What is the bias-variance trade-off? How to address bias and variance respectively?

2) [5 pts] What is overfitting? List several common techniques that can reduce overfitting.

3) [5 pts] What is learning rate? Why learning rate cannot be too large or too small.

2. use the following confusion matrix of a classifier:

		Predicted results	
		True	False
Actual values	True	80	30
	False	40	70

1) [5 pts] What is the precision? Please show your calculation.

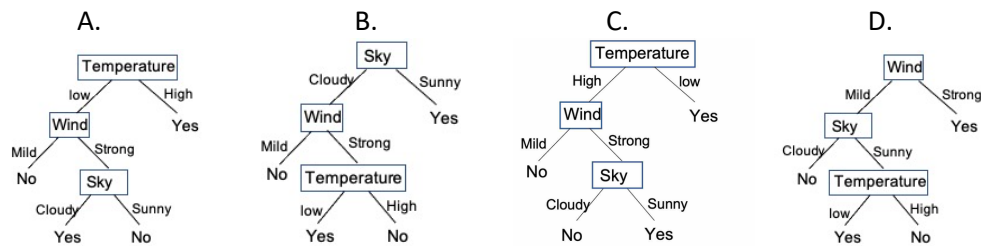
2) [5 pts] What is the recall? Please show your calculation.

3) [5 pts] What is its F_1 -score? Please show your calculation.

3. Consider training a decision tree using the training set S as follows:

Sky	Temperature	Wind	EnjoySport
Cloudy	Low	Mild	Yes
Cloudy	High	Strong	No
Sunny	High	Strong	Yes
Sunny	High	Mild	No

- 1) [4 pts] What is the sample entropy of the training set S?
- 2) [5 pts] What is the information gain of attribute Sky over S?
- 3) [5 pts] What is the information gain of attribute Temperature over S?
- 4) [6 pts] Which of the following is the decision tree we will obtain using the training set S?



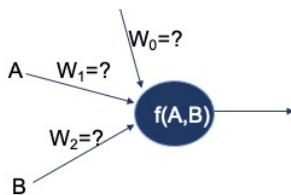
- 5) [4 pts] What is the predicted result for below example using the decision tree in 4)?
(Sky = Cloudy, Temperature = Low, Wind = Strong)
- 6) [5 pts] What is the accuracy rate of the decision tree obtained in 4) over training set S?
- 7) [6 pts] Can the tree in 4) be further pruned without losing accuracy? Why?

4. [6 pts] Use the Naïve Bayes Algorithm to predict a new instance based on a dataset with 10 examples below. Show your calculation.

Day	Outlook	Temperature	Humidity	Wind	PlayTennis
D1	Sunny	Hot	High	Weak	No
D2	Sunny	Hot	High	Strong	No
D3	Overcast	Hot	High	Weak	Yes
D4	Rain	Mild	High	Weak	Yes
D5	Rain	Cool	Normal	Weak	Yes
D6	Rain	Cool	Normal	Strong	No
D7	Overcast	Cool	Normal	Strong	Yes
D8	Sunny	Mild	High	Weak	No
D9	Sunny	Cool	Normal	Weak	Yes
D10	Rain	Mild	Normal	Weak	Yes

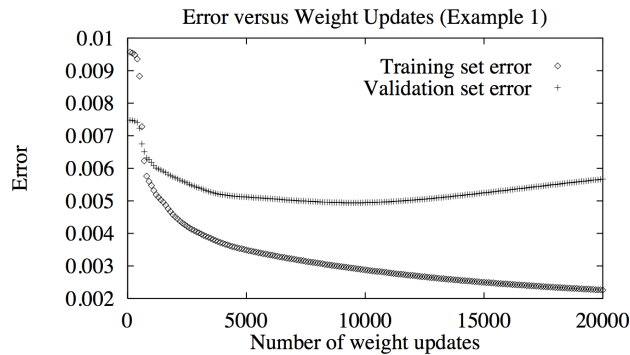
The new instance is: $\langle \text{Outlook}=\text{Rain}, \text{Temperature}=\text{Mild}, \text{Humidity}=\text{High}, \text{Wind}=\text{Weak} \rangle$

5. Consider the following two-input perceptron. Now we want to choose appropriate weights of the inputs to turn the perceptron into a Boolean function $f(A, B)$.



- 1) [8 pts] Decide weights (W_0 , W_1 and W_2) that make $f(A, B) = A \text{ OR } \neg B$. Please show your calculation.
- 2) [5 pts] Which of the following the Boolean function $f(A, B)$ can NOT represent?
 - A. AND
 - B. OR
 - C. XOR
 - D. NAND

6. Consider below training error and test error observed as we train for a neural network using batch gradient descent.



- 1) [3 pts] Is there overfitting with the trained model? How do you know?
- 2) [3 pts] i) If we double the size of the training data, plot the new curves (on the figure together with the old curves) for training error and testing error, respectively.
[5 pts] ii) Briefly explain why we may have such new curves.

7. Consider three classifiers that make the following predictions for a sample X. We want to make the final decision using the ensemble of the three classifiers.

Sample x	Result
Classifier 1	Class 2
Classifier 2	Class 2
Classifier 3	Class 1

- 1) [3 pts] If we use the majority vote as the fusion function, what is our final decision?
- 2) [5 pts] If we use the weighted majority vote as the fusion function, what is our final decision?
Given the following weights of the classifiers.
Classifier 1: 0.2
Classifier 2: 0.6
Classifier 3: 0.2

- 3) [7 pts] If we use the Naïve Bayes method as the fusion function, what is our final decision? Given the following confusion matrix of the classifiers. Show your calculation.

i) Classifier 1

	Class1	Class2
Class1	70	10
Class2	30	50

ii) Classifier 2

	Class1	Class2
Class1	80	30
Class2	20	70

iii) Classifier 3

	Class1	Class2
Class1	80	10
Class2	40	40