

KATHMANDU UNIVERSITY
End Semester Examination [C]
December, 2018

Marks scored:

Level : B. Sc.
Year : IV

Course : COMP 484
Semester: I

Exam Roll No. :

Time: 30 mins.

F. M. : 10

Registration No.:

Date **DEC 30 2018**

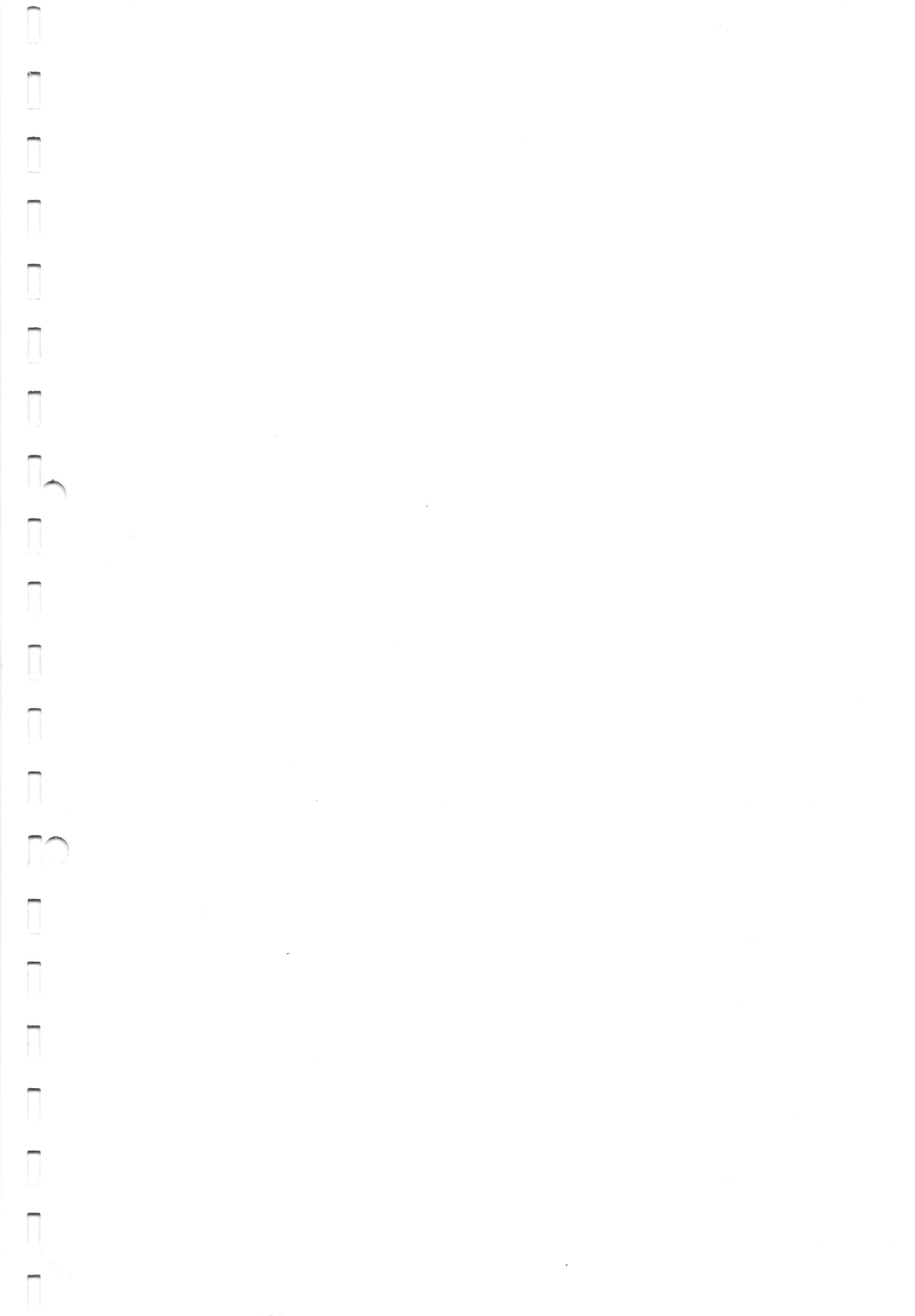
SECTION "A"
[20 Q. × 0.5 = 10 marks]

Encircle the most appropriate answer.

1. The correlation between the number of years an employee has worked for a company and the salary of the employee is 0.75. What can be said about employee salary and years worked?
 - a. There is no relationship between salary and years worked
 - b. Individuals that have worked for the company the longest have higher salaries.
 - c. Individuals that have worked for the company the longest have lower salaries.
 - d. The majority of the employees have been with the company a long time.
2. Bootstrapping allows us to:
 - a. Choose the same training instance several times
 - b. Choose the same test set instance several times
 - c. Build models with alternative subsets of the training data several times
 - d. Test a model with alternative subsets of the test data several times
3. Which one of the following is a categorical value?
 - a. RMSE
 - b. RSquared
 - c. Accuracy
 - d. All of the mentioned
4. The average positive difference between the computed and desired outcome values is:
 - a. Root mean squared error
 - b. Mean squared error
 - c. Mean absolute error
 - d. Mean positive error
5. Selecting data so as to assure that each class is properly represented in both the training and test set:
 - a. Cross validation
 - b. Stratification
 - c. Verification
 - d. Bootstrapping
6. Data used to optimize the parameter settings of a supervised learner model:
 - a. training
 - b. test
 - c. verification
 - d. validation
7. Logistic regression is a _____ regression technique that is used to model data having a _____ outcome.
 - a. Linear, numeric
 - b. Linear, binary
 - c. Non-linear, numeric
 - d. Non-linear, binary
8. Machine learning techniques differ from statistical techniques in that machine learning methods
 - a. Typically assume an underlying distribution for the data
 - b. Are better at dealing with missing and noisy data
 - c. Are not able to explain their behavior
 - d. Have trouble with large-sized datasets

9. Suppose your model is overfitting. Which of the following is NOT a valid way to try to reduce overfitting?
 - a. Increase the amount of training data
 - b. Improve the optimization algorithm being used for error minimization
 - c. Decrease the model complexity
 - d. Reduce the noise in the training data
10. Higher value of entropy would mean which of the following:
 - a. Attributes of instances belong to a mixed class of values
 - b. Attributes of instances belong to just one class of values
 - c. Attributes of instances cannot be related to any class of values
 - d. None of the above
11. Regarding bias and variance, which of the following is true? (Here “high” and “low” are relative to the ideal model.)
 - a. Models which overfit have a high bias
 - b. Models which overfit have a low bias
 - c. Models which underfit have a high variance
 - d. None of the above
12. The MDL principle favors:
 - a. Simpler hypotheses
 - b. Complex hypotheses
 - c. A mix of simple and complex hypotheses
 - d. None of the above
13. A “Bayesian Network” is most accurately described as:
 - a. A special case of neural network that makes use of the Bayes’ Theorem
 - b. The network variant of Bayes Theorem, assuming independent features
 - c. A probabilistic model of which Naïve Bayes is a special case
 - d. A network of probabilistic learning functions, connected by Bayes Rule
14. Averaging the output of multiple decision trees helps:
 - a. Increase bias
 - b. Increase variance
 - c. Decrease bias
 - d. Decrease variance
15. What is true about the delta training rule in Artificial Neural Networks?
 - a. It handles well with non-linearly separable training examples
 - b. It converges only asymptotically toward minimum error hypothesis
 - c. Both “a” and “b”
 - d. It converges after a finite number of iterations to a hypothesis that perfectly classifies the training data
16. Which of the following algorithm deals with concept learning?
 - a. ID3
 - b. Naïve Bayes
 - c. Candidate Elimination
 - d. Backpropagation

17. Supervised learning and unsupervised clustering both require at least one :
- a. Hidden attribute
 - b. Output attribute
 - c. Input attribute
 - d. Categorical attribute
18. The process of forming general concept definitions from examples of concepts to be learned:
- a. Deduction
 - b. Abduction
 - c. Induction
 - d. Conjunction
19. With Bayes classifier, missing data items are :
- a. Treated as equal compares
 - b. Treated as unequal compares
 - c. Replaced with a default value
 - d. Ignored
20. Which of the following does not deal with training the model with labeled examples?
- a. Decision Trees
 - b. Concept Learning
 - c. Instance-based Learning
 - d. None of the mentioned



KATHMANDU UNIVERSITY
End Semester Examination [C]
December, 2018

DEC 30 2018

Level : B.Sc.
Year : IV
Time : 2 hrs. 30 mins.

Course : COMP 484
Semester: I
F. M. : 40

SECTION "B"

[6 Q. × 4 = 24 marks]

Attempt *ANY SIX* questions.

1. Distinguish between supervised and unsupervised learning methods. Provide an example for each of them.
2. Define "bias" and "variance". What is the trade-off between these two?
3. What is overfitting? How do you avoid overfitting?
4. What is the primary difference between "gradient descent" and "stochastic gradient" algorithms? When would you use "gradient descent" over "stochastic gradient descent" and vice-versa?
5. What is cross-validation? Give examples of cross-validation methods.
6. Explain "prior hypothesis" and "posterior hypothesis". What is a "maximum a posteriori probability"?
7. What are the advantages and disadvantages of "neural networks" as a Machine Learning classifier?

SECTION "C"

[2 Q. × 8 = 16 marks]

Attempt *ANY TWO* questions.

8. Discuss the Backpropagation algorithm. What are its concrete advantages over single perceptrons? Also point its drawbacks. [4+ 2 + 2]
9. How does a decision tree learning algorithm work? What is the basis for selection of the nodes of the decision tree? What types of problems are suitable to tackle using decision trees? [3 + 2 + 3]
10. What is the key idea in the CANDIDATE-ELIMINATION algorithm? Define the term "version space". Explain the most specific and general hypotheses. [3 + 3 + 2]

