Final Exam - Fall 2023

Started: Dec 15 at 4:28am

Quiz Instructions

Dear students,

This is an open book, open notes exam. Open Canvas course materials.

Internet search is not allowed.

Please answer all questions. You have three (3) hours (180 minutes) so that you can answer in a relaxed manner. You may start 15 minutes early and complete 15 minutes late, hence you have 210 minutes total.

There are Fill-in-the Blanks, and Match the Following questions and some problem-solving questions, however, each question has a little more weight in points. The test has been designed so that you can complete answering all questions in less than two hours.

For the questions that requires problem-solving, you can work the problem out in clear handwriting or electronically in a Word document (please convert to PDF). If you prefer to minimize the number of attachments, instead of submitting one document per question, it is okay if you submit one single document for all questions. You can submit an attachment as an answer directly to the question or include the answers to all questions in a single attachment to the last question that says: "Submitting one document for all problems? Submit here". I will be able to see that all the answers are in a single document. It is preferable for grading convenience, that you make an attachment if required, directly to the question itself, instead of providing answers to all problem-solving questions in a single document. Please remember, the problem-solving questions are very simple in nature. They are intended only to test your understanding of the basic concept.

All the questions must be answered individually. Collaboration or sharing solutions or seeking solutions or searching on the internet is strictly NOT ALLOWED as this is an exam. Use of tools such as Chat GPT is strictly prohibited. Any violation will be viewed as a violation of the university honor code. So please keep this in mind while answering the exam.

There is only one attempt, meaning you can hit the submit button only once. CANVAS saves your answers each time you answer a question. You can exit without hitting the submit button and come back in later and answer the remaining questions or change your answers to questions as long as you have not hit the submit button. However, once you begin the test, you have only 210 minutes total to complete.

If any student hits the submit button accidentally before you complete the exam, a request to give another chance will not be honored. Hence, be extremely careful.

Best wishes and cheers,

:)

Jagan

Question 1 — 3 pts

Bias is associated with _______________ as opposed to overfitting.

Question 2 — 3 pts

Deep learning uses mostly the --------- activation function in the hidden layers.

Question 3 — 3 pts

K-means has problems (or issues) when the data contains _________.

Question 4 — 3 pts

Sampling is the main technique for ______________ the size of data from the population size to a manageable size.

Question 5 — 3 pts

A neural net may as well be accomplished with just one ---------------.

Question 6 — 10 pts

Match the following approaches based on centroid-based clustering using K-Means

Centroid-based clustering	[Choose]
Density-based clustering	[Choose]
Distribution-based Clustering	[Choose]
Hierarchical Clustering	[Choose]
K - Means Clustering	[Choose]

Question 7 — 5 pts

Which of the following evaluation metrics cannot be applied in case of logistic regression output to compare with target?

- Mean-Squared-Error
- AUC-ROC
- Accuracy
- LogLoss

Question 8 — 5 pts

What are the characteristics of an Artificial Neural Network (ANN)?

- It can handle redundant and irrelevant attributes because weights are automatically learnt for all attributes
- It is sensitive to noise in training data
- It represents a hierarchy of features at multiple levels of abstractions
- It is extremely fast in handling missing attributes

Question 9 — 5 pts

In k-NN what will happen when you increase/decrease the value of k?

- The boundary becomes smoother with decreasing value of K
- Smoothness of boundary doesn't dependent on value of K
- The boundary becomes smoother with increasing value of K
- None of these

Question 10 — 5 pts

Which of the following is an example of predicting regression?

- Predicting wind velocities
- Predicting sales amounts of new product based on advetising expenditure.
- Time series prediction of stock market indices
- Predicting tumor cells as benign or malignant

Question 11 — 5 pts

Think on your feet!:

Explain in your own words, how to find the optimal value of K in KNN? What if the value of K is too large or too small?

p — 0 words

Question 12 — 5 pts

Explain the following (think on your feet!):

In the neural net topic, the back propagation of error is used iteratively in a cycle to update the weight and other parameters. a

When does the updating stop? The most common conditions are one of the following:

When the new values of the bias and weights are only incrementally different from those of the preceding iteration
When the misclassification rate reaches a required threshold
When the limit on the number of runs is reached

Please explain the above three conditions with your own intuition what they mean and why these conditions may exist.

p — 0 words

Question 13 — 5 pts

Why is logit called a regression? Does it simulate a line even though data is not a line? (Explain in detail)

p — 0 words

Question 14 — 5 pts

Explain SVM and what are slack variables in SVM?

p — 0 words

Question 15 — 15 pts

Explain Entropy and Information Gain and calculate the entropy and information gain for the following.

Problem:

Given two dice (each with six numbers from 1 to 6):

(a) what is the entropy of the event of getting a total of greater than or equal to 11 in one throw?

(b) what is the entropy of the event of getting a total of equal to 8 in one throw?

What is the Information GAIN going from state (a) to state (b)? Explain what this means in your own words.

Please show all working and the final result. Show all working. You may do it on paper and upload your working (if you are using a cell phone camera app, make sure the writing on each page is very clear and you upload a PDF file). Details of working are as important as your final answer.

p — 0 words

Question 16 — 20 pts

Problem Solving:

The following set of records have the following classes:

Answer the question below with respect to the following table. Provide clear steps.

Height	Weight	Class
low	low	no
high	low	yes
high	medium	no
medium	medium	no
low	high	yes
medium	high	yes

Explain how correction is applied when probabilities are zero to get non-zero probabilities in Naive Bayes classification. (5 points)

Using Naive Bayes method, find the class for the test record with height 'medium' and weight "high".

(a) with the traditional naive bayes independence condition, and (6 points)

(b) without the independence condition (6 points)

Compare and explain the reasons for the differences and what the differences mean intuitively. (3 points)

Show all working. You may do your working on paper and upload your working (if you are using a cell phone camera app, make sure the writing on each page is very clear and you upload a PDF file). Details of working are as important as your final answer.

Edit View Insert Format Tools Table
12pt Paragraph B I U

p — 0 words

Question 17 — 0 pts

Upload Choose a File

Quiz saved at 4:29am — Submit Quiz

Questions
Question 1
Question 2
Question 3
Question 4
Question 5
Question 6
Question 7
Question 8

Time Running: Hide Time
Attempt due: Dec 16 at 11:59pm
3 Hours, 29 Minutes, 43 Seconds