CS 103 Autumn 2022: End-semester Exam Solutions

Name:_____ **Roll number:**_____

1. AlphaGo was a program that defeated the human champion in the game of Go in 2016. This program was trained in two phases. In Phase 1, a neural network model was trained on a data set D to imitate the decision making of expert human players. D contained a set of boards encountered in the games played by experts, and for each board, included the move played by the expert. The model was trained to predict the board-to-move associations in D. In Phase 2, the trained model from Phase 1 was made to play itself (in the sense that both players would invoke the same model for making their moves). Based on the outcome of the each such "self play" game, the model would be slightly tweaked to encourage the winning sequence of moves and to discourage the losing sequence of moves. Which among the two phases is an example of supervised learning? Answer with suitable justification.

   Ans: In supervised learning, the training data set contains input-output pairs (x_i, y_i). A model trained on this data can predict the output y for an unseen input x. Phase 1 in the training of Alpha precisely matches this description: each x_i is a board and the corresponding label y_i the expert move.

2. We discussed these popular AI programs in class: IBM Watson, Dall-E, ChatGPT, Waymo, Alexa. Describe any two of these programs in 1-2 lines.

   Ans: Watson was a program built by IBM to participate in the Jeopardy quiz competition. Using knowledge representation and reasoning, the program was engineered for question-answering. It went on to win the Jeopardy competition in 2011.

   Dall-E, created by Open AI, uses a large language model to generate images from text phrases.

   ChatGPT, also created by Open AI, uses a large language model in its interactions (chat) with a user. It is able to write poems, essays, generate jokes, and so on. It become especially popular around December-2022.

   Waymo is a company that makes self-driving cars, some of which are in deployment as taxis in particular regions of the United States. The company is now a part of Google.

   Alexa is a hardware device the interacts with users through speech. It performs speech-to-text to translate their audio queries into natural language, then retrieves response from a database, and communicates them again through speech.

3. Consider a machine learning model for the salary of a working professional (say, in rupees) as a function of the person's education and work experience (measured in, say, years), using the equation: salary = W_0 + W_1 * education + W_2 * experience + W_3 * education * experience. The weights W_0, W_1, W_2 and W_3 are estimated by the model as those that minimize error on training data obtained from many professionals. Once the weights are computed, the model can be used to predict the salary of a professional given his/her education and experience.

   (a) Is this model an example of linear regression? Answer yes/no with justification.

   Ans: Yes, this is a linear regression model, because it is linear in the weights.

(b) Give an example where you think the predictions from the model can go wrong in real life.
Ans: Many possible answers exist. For example, people may get lucky and find high-paying jobs that pay a lot more than the education or experience justifies.

4. The Wikipedia contains about 40 million English pages with about 6 billion words. Let count(X,Y) denote the number of times a word X appears in the *context* of (e.g., in the same sentence as) another word Y. Which of the following is the most likely approximation of the probability that the word 'bark' appears in the context of the word 'dog'? In the options below, N is the total number of words in the vocabulary, and 'v' denotes any word in the vocabulary

   (a) count(bark, dog) / N
   (b) count(bark, dog) / sum of count(v, dog) for all v
   (c) count(bark, dog) / N*N
   (d) maximum of count (bark, v) / count (v, dog) over all v

   Ans: (b)

5. Which of the following CANNOT CONVINCINGLY be used as the evidence that all languages originated from one single language? Tick all the correct options and none of the wrong ones.

   (a) Sanskrit and Latin for 'eight' are 'ashta' and 'octo' respectively
   (b) About 90% of world's languages place the subject before the object
   (c) 'procrastination' and 'delay' mean the same thing
   (d) Malayalam belongs to the Dravidian family of languages

   Ans: (c) and (d)

6. You are given "parallel" sentences between Czeck and English as follows:

   - *nesu* "I carry"
   - *ponese* "He will carry"
   - *nese* "He carries"
   - *nesou* "They carry"
   - *yedu* "I drive"
   - *plavou* "They swim"

   In Czeck, the word for 'but' as a conjunction linking two sentences is 'ale'. Which of the following is the correct meaning of the Czeck sentence "poplavou ale poyedu"?

   (a) They carry but I swim
   (b) He will swim but I will drive
   (c) They will swim but I will drive
   (d) He will carry but I will swim

   Ans: (c)

7. Consider the two-player games shown below. Each game has two strategies A and B for player 1, and two strategies P and Q for player 2. For each combination of strategies chosen by the players, the payoffs or rewards for the two players are shown as a tuple against the strategy combination in the tables below. For example, for the first game shown in part (a) below, if the first player chooses A and the second player chooses P, then the payoff is 2 for the first player and 3 for the second, as shown by the entry "2,3" in the table below. For each of the games below, does a dominant strategy equilibrium exist for the players? Ans yes/no. If yes, you must also state the equilibrium strategy. Recall that a dominant strategy equilibrium is one in which both players have a dominant strategy irrespective of what the other player chooses.

(a)

Game (a)	P	Q
A	2,3	1,5
B	0,0	4,1

(b)

Game (b)	P	Q
A	7,2	4,5
B	6,0	2,2

Ans: No dominant strategy exists for game (a). For game (b), it is (A,Q).

8. For each of the games in the previous question, answer if a Nash equilibrium exists. If you answer yes, state the combination of strategies in the equilibrium as well. Recall that a Nash equilibrium occurs when either player has no incentive to change their strategy from the equilibrium choices.

   Ans: For game (a), it is (B, Q). For game (b), it is (A, Q).

9. Below are the events that occur in modern computers when a user writes, compiles and executes a program, presented in a jumbled order. State the correct chronological order in which the events occur, from earliest to latest. You can write your answer by listing the statement numbers in order, e.g., "(e), (d), (c), (b), (a)"

   (a) CPU fetches instruction and data from main memory.

   (b) User writes program in an editor and stores the source code file on disk.

   (c) The compiled instructions and data are loaded into main memory from disk.

   (d) User compiles the program, compiler translates the high-level code into machine code.

   (e) CPU executes instruction and stores any generated result into memory.

   Ans: b, d, c, a, e

10. Consider the design of a simplified toy CPU. Suppose every instruction is processed by the CPU in three stages: fetching the instruction from memory, decoding the instruction, and executing it. Suppose each of the three stages takes 1 millisecond (10^{-3} seconds) each to run.

    (a) Suppose the CPU is designed in a non-pipelined manner. That is, the next instruction is processed by the CPU only after the previous instruction has completed all three stages.

In this design, what is the time taken to execute a single instruction? Further, how many instructions per second can the CPU complete?

Ans: 3 millisec per instruction, 333 instructions per second.

(b) Now suppose the CPU uses a pipelined design instead, with the three stages working in parallel. That is, when an instruction is in the execute stage, the next instruction will be in the decode stage, and the instruction after it will be in the fetch stage of the pipeline. In this pipelined design, what is the time taken to execute a single instruction? Further, how many instructions per second can the CPU complete?

Ans: 3 millisec per instruction, 1000 instructions per second in steady state (998 in first second).

11. Suppose you are managing a manufacturing company where the users use several software applications for their day-to-day work. Which of the following is/are some of the benefits of accessing these user software applications as Software-as-a-Service on the cloud over the Internet (e.g., like Google Docs), as compared to deploying the software on computer hardware setup locally within the company? You must tick all the correct options and none of the wrong ones.

    (a) Applications can be accessed with lower network latency/delay.

    (b) No need to manage the power and cooling of the hardware on which the application runs.

    (c) No need to worry about the installation and updates to the application software.

    (d) No need to train users to use the application software.

    Ans: (b), (c)

12. The following are events (listed in a jumbled order) that occur when a user browses a simple web page using a web browser. State the correct chronological order in which the events occur, from earliest to latest. You can write your answer by listing the statement numbers in order, e.g., "(e), (d), (c), (b), (a)"

    (a) User types the name of the website in the address bar of the browser.

    (b) The web browser displays the HTML webpage to the user.

    (c) The web server sends a HTML webpage as a response back to the browser.

    (d) DNS is used to resolve the website name into the IP address of the web server.

    (e) User's browser sends a HTTP GET request to the web server.

    Ans: (a), (d), (e), (c), (b)

13. Consider the communication between two nodes A and B connected by a copper wire. A wishes to send a file with 250 characters to B. Each character is represented by 8 bits in the computer. During transmission over the wire, bits 0 and 1 are represented as low and high voltages. The sender maintains each bit's voltage signal on the wire for 1 millisecond, within which time the receiver will decode the signal back to the bit. You can assume that there are no errors in the transmission.

    (a) What is bandwidth (in bits per second) of the communication link between A and B?
    Ans: 1000 bits per second

(b) How long does it take for A to send the file to B?
Ans: 2 seconds

14. The Internet uses the Transmission Control Protocol (TCP) to transfer data packets between sending and receiving computers reliably and in order (i.e., in the same order as sent) over network links which can drop (i.e., lose) or reorder (i.e., jumble) packets. Consider the following description of a simplified version of TCP between a sender A and receiver B. A wishes to send some packets to B using TCP. To keep track of which packets have reached B and which have not, A attaches a sequence number to every packet. Receiver B sorts the received packets by their sequence numbers and consumes them in the order of their sequence numbers. When B receives a packet from A, it sends back an acknowledgement packet (lets call it ack) which has the sequence number of the last packet that B has consumed. For example, suppose B has received a packet with sequence number N, and all other packets before it as well. Since all packets are in order, B can consume all of them, and it sends back an ack with sequence number N. Now, suppose packet $N + 1$ is lost in the network, and B receives packet number $N + 2$ next. On receiving this packet, it still sends back an ack with number N, because B cannot consume packet $N + 2$ until it has received packet $N + 1$. So, when packet loss occurs (e.g., packet number $N + 1$ is lost and $N + 2$ is received at B), A will see a duplicate ack (i.e., a second ack with the same number N). When A gets 3 such duplicate acks (say, when packets $N + 2$, $N + 3$ and $N + 4$ were received at B), then A uses this to guess that packet $N + 1$ was lost and proceeds to retransmit (resend) it on the network. Suppose A sends packets with sequence numbers 0 to 7 to B in this order. Of these, suppose packet with sequence number 2 is lost in the network. All other packets are received at B in the order they were sent in. Answer the following questions using the description above.

(a) State the sequence numbers of the acks sent by B for the first five packets it receives.
Ans: 0, 1, 1, 1, 1

(b) A decides it must retransmit the lost packet number 2 only after B receives packet number X. What is X?
Ans: Packet number 5 (A receives 3 duplicate acks when packets 3, 4, 5 are received at B).

15. Consider the following recursive algorithm (also called a divide-and-conquer algorithm) to compute the convex hull of a set of N points in a plane. Sort the points from left to right, and divide them into two halves of $N/2$ points each (you can assume N to be even for now). Next, recursively construct a convex hull over the left and right halves of $N/2$ points. Next, merge the left and right convex hulls by joining them using two "bridges" (line segments joining the two convex hulls), one at the top and one at the bottom. Some details of the algorithm (e.g., how do you merge the two convex hulls, or how do you find the bridges) are left unspecified for you to guess. Using this rough description, draw two diagrams next to each other in the space provided below. In the first diagram, put down 12 points in a plane, and draw the left and right convex hulls for the two sets of 6 points each. (You can visually figure out the convex hull for the 6 points, and there is no need to show another iteration of the recursion.) In the second diagram next to it, show how the two convex hulls can be merged using bridges, and show the final convex hull.

Ans: Many possible correct answers exist.

16. Recall the Big-O notation for the computational complexity of an algorithm. For example, an algorithm whose run time is a linear function of number of inputs n is said to have a complexity

of O(n). An algorithm whose run time is a constant and does not depend on number of inputs is said to have a complexity of O(1). Consider the following problem: given an unsorted array of n integers, in which multiple elements can be equal to each other, we wish to find *any* array element that is not equal to the minimum element (assuming it exists). Consider the following algorithm to find such an element: go over the elements in the array one by one until two unequal elements are found, and return the higher of the two (since it is guaranteed to not be the minimum).

(a) What is the complexity of the algorithm in the best case (i.e., for the input arrays which result in the most favorable execution time)?
Ans: O(1)

(b) What is the complexity of this algorithm in the worst case?
Ans: O(n)

17. Consider the following simple program to find the minimum element of an array "A" of size N, written in C-like code.

```
int min = INFINITY; //some large value
int i = 0;
while (i < N) {
  if(A[i] < min)
    min = A[i];
  i = i+1;
}
```

State the loop invariant for the while loop in this program (in English or in mathematical notation).

Ans: min is the minimum element of the first i elements of the array.

18. Consider the following class scheduling problem, which is computationally hard. A college has N classes, which have to be scheduled in M time slots, subject to various scheduling constraints around the availability of rooms and curriculum of students. We wish to obtain a solution to this problem by modeling it using boolean constraints and feeding the constraints to a SAT solver, as seen in class. Suppose we use boolean variables $x_{ij}, i \in 1..N, j \in 1..M$ where x_{ij} is true if class i is scheduled in slot j. Write down constraints on x_{ij} which capture the following:

(a) Any class i can be scheduled in only one of the M slots.
Ans: For each $i \in 1..N$, exactly one of $x_{i1}, x_{i2}, ...x_{iM}$ is true.

(b) Two specific classes p and q cannot be scheduled in the same slot, for some $p, q \in 1..N$.
Ans: For all $j \in 1..M$, NOT x_{pj} OR NOT x_{qj}. In other words, NOT (x_{pj} AND x_{qj})

19. Consider the 3-color graph coloring problem discussed in class, where we wish to color the vertices of a graph with 3 colors in such a way that no two vertices connected by an edge have the same color. Draw below an example of a graph that can be colored with 3 colors, and an example of a graph that cannot be colored in 3 colors. Clearly show the vertices, edges, and colors (C1, C2, C3) assigned to the vertices in your drawing.

Ans: Many possible correct answers exist.

20. Suppose Alice and Bob wish to securely exchange a message *m* with each other, without an eavesdropper Eve (who listens to all data exchanged between them) learning anything about the message. For this communication, Alice and Bob agree on a shared secret one-time pad *b* and Alice sends ciphertext $c = m \oplus b$ to Bob. Which of the following statements is/are true? Tick all the correct answers and none of the wrong answers.

    (a) Bob recovers the original message by computing $c \oplus b$

    (b) Bob recovers the original message by computing $m \oplus b$

    (c) Eve can learn of the original message if she knows *b*

    (d) Eve can learn of the original message if she knows *c*

    Ans: (a), (c)

21. State if each of the statements below is falsifiable or not. If you answer falsifiable, you must also state how you can potentially falsify it. Otherwise, you must explain why the statement is not falsifiable.

    (a) Drinking a juice made with the hair of giraffes will make me grow taller by 1 cm every day.
    Ans: Falsifiable, by drinking the juice and measuring if one grows tall.

    (b) All crows are black.
    Ans: Falsifiable, by looking for a crow of another color.

22. The following arguments have an implied premise that must be assumed in order for the arguments to be cogent. State the implied premise in each case.

    (a) My friend broke his leg. So he cannot join us tomorrow for the hike up the Sameer Hill.
    Ans: One cannot climb up the hill with a broken leg.

    (b) I just heard the doorbell ring. There is someone at the door.
    Ans: Someone must be at the door for the bell to ring.

23. Give one example each of a weak inductive argument, having the following fallacies:

    (a) Improper analogy in analogical induction
    Ans: Many correct answers exist. Many of you seem to have gotten this question wrong. Note that analogical induction is not just analogy, but actually generalizing a property about the whole population based on an analogy. For example, drug X cures disease Y in rats. So it will do so in humans as well.

    (b) Biased sampling
    Ans: Many correct answers exist.

24. Give one example each of an argument that uses an invalid deduction to arrive at the conclusion, specifically:

    (a) Denying the antecedent (i.e., premises of the type "if A then B" and "not A")
    Ans: Many correct answers exist.

    (b) Affirming the consequent (i.e., premises of the type "if A then B" and "B is true")
    Ans: Many correct answers exist.

25. Give one example each of an argument with the following fallacy.

    (a) False equivalence (comparing apples and oranges)
    Ans: Many correct answers exist.

    (b) False dilemma (either-or fallacy)
    Ans: Many correct answers exist.

26. Consider the following argument by the prosecutor in a murder investigation that occurred on a busy main road in Mumbai, a city with approximately 20 million people. "The DNA evidence found at the crime scene matched that of person X. If X is indeed the murderer, the probability of finding his DNA on the crime scene is very high, close to 100%. On the other hand, the chance that a person's DNA matches the DNA found at the crime scene if he/she wasn't the murderer (and therefore not present at the crime scene) is very low, almost one in a million. Therefore, X should be convicted for the murder as the probability that he is guilty is very high." For each of the scenarios below, state whether the above argument is cogent or not, in the context of the additional information provided below. For each case, you must also provide a suitable justification for your answer, stated in terms of conditional probabilities.

    (a) The DNA at the crime scene was matched with DNA samples in the master database of the city, which includes samples collected from a large number of people for various purposes across many years. X is one of these randomly sampled people.
    Ans: Consider the following example numbers. Suppose all 20 million people were tested. About 20 innocent people, in additon to the actual murderer, may get a false positive DNA match.

	DNA match	DNA not match
Guilt	1	0
Not guilty	20	20 million - 20

    From the above table, probability of DNA match given someone is guilty is very high. But probability that someone is guilty given DNA match is only 1 in 21. This probability is too low to convict. Hence this is a fallacious argument. (This fallacy is commonly referred to as the prosecutor's fallacy.)

    (b) The DNA at the crime scene is matched with DNA samples collected from ten suspects who have a strong motive to harm the murder victim. X is one of these ten suspects.
    Ans: We can argue similarly as above, except that the pool of people being tested now is not all 20 million but only ten suspects. So the probability of a false positive DNA match is very low. Therefore, the probability that someone is guilty if DNA has matched is pretty high. Hence this is a cogent argument.

    Note the marking scheme for above question. Fully correct answers that provided justification in terms of conditional probabilities got full 2 marks. Other answers which were correct but did not explain in terms of conditional probabilities got 1 mark. Answers which were partly correct (e.g., guessed cogent for both parts) and without proper justification did not get any marks.

27. We have seen several examples of "vernacular" science in class, which was argued as being different from "frontier" science that is often done in top science and technology institutes of the world.

In your opinion, should an institute like IIT Bombay (that is located in a developing country) work on problems that the bottom of the society faces, which may not be of interest to scientists pursuing frontier science? Answer honestly (there is no one right answer we are looking for) and provide a suitable justification to support your opinion.

Ans: Any sensible answer was given marks.

28. We have seen several examples of computer science projects that were being developed or deployed for "social good". Briefly describe (in 1-2 sentences) an example from your experience where a computer science idea has been used for social good. Describe the idea as well as its impact. You are expected to provide an example that has NOT been discussed in class.

    Ans: Any sensible answer was given marks.

29. The course discussed how one should scientifically analyze real life problems using data to arrive at sensible conclusions. Consider the problem of infrastructure spending by the government on constructing roads, metros, rail networks and so on. Given the large amounts of money spent, one would like to know if the infrastructure made any noticeable impact to the public. One way to do this is as follows. Suppose one wishes to evaluate the impact of a new infrastructure project in a city. Pick N important points in the city. Compute the minimum time taken to travel between every pair of points at several times during the day, by looking up some service like Google Maps, which provides the best possible time across all modes of transport (public and private). Calculate this time across several days and at different times of the day, and take an average of all these samples. Compare this average before and (and a suitable amount of time) after the infrastructure is constructed and put in use. If the average travel time across city landmarks improves significantly, then one can say that building the infrastructure was a good idea.

    (a) Do you think this is a good metric? For example, is it easy to compute? Is it useful to assess the impact of infrastructure projects? Answer with suitable justification.

        Ans: Any sensible answer was given marks.

    (b) Suggest another metric (different from the one above) that can be used to assess the impact of infrastructure projects.

        Ans: Any sensible answer was given marks.

30. Give one suggestion to improve this course in its next offering. For example, you can suggest some topic that could be taught or skipped, or suggest changes to the way the course is run.

    Ans: Any useful feedback was given marks.