NAME: _____

AI2 Midterm Fall 2025

Each question is worth 10 points.

1. A robot moves on a grid where 20% of cells are dangerous. Its camera sensor detects danger correctly 90% of the time but falsely signals danger in 10% of safe cells. If the camera sensor signals danger, what is the probability the cell is actually dangerous?

2. Continuing from last question, the robot also has IR camera sensor. The IR camera sensor detects danger correctly 80% of the time but falsely signals danger in 20% of safe cells. If the camera *and* IR camera sensors both signal danger, what is the probability the cell is actually dangerous?

3. Continuing from last question, the robot also has a tactile sensor that is independent of both camera sensors. The tactile sensor detects danger correctly 60% of the time but falsely signals danger in 30% of safe cells. If the camera *and* tactile sensors both signal danger, what is the probability the cell is actually dangerous?

4. Consider a 2-layer neural network with 2 inputs, 2 neurons in the hidden layer, and 1 output neuron. The network uses the ReLU activation function, defined as $\text{ReLU}(x) = \max(0, x)$. The network weights and biases are given as:

$$\text{Hidden layer weights: } W_1 = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}, \quad b_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

$$\text{Output layer weights: } W_2 = \begin{bmatrix} 1 & 2 \end{bmatrix}, \quad b_2 = 3$$

The input to the network is:

$$x = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

Calculate the final output of the network. Show work (calculate the output of the hidden layer, apply ReLU, etc.)

5. Suppose we wish to minimize the output of the network. What would the gradients be for W_1, b_1, W_2, b_2?

6. From the below table, which attribute is most relevant for identifying fraudulent transactions. Explain your reasoning (Hint: you do not need to calculate information gain).

Transaction	Amount > $1000	International	Unusual Time	Fraud?
T1	Yes	No	No	No
T2	Yes	Yes	Yes	Yes
T3	No	No	Yes	No
T4	Yes	Yes	No	Yes
T5	No	Yes	Yes	No
T6	Yes	No	Yes	Yes

7. Continuing from last question, draw a decision tree to identify fraudulent transactions.

8. You are building a checkers playing program. You need to explore possible moves. You have a choice between depth-first and breadth-first search methods. Which one is appropriate. Explain why.

9. We train a model to predict fradulent transactions based on 10000 different features collected from everywhere. After training, the model achieves an accuracy of 98% on the training set. We put the model into production, and discover it performs poorly at actually predicting fradulent transactions. What went wrong? Provide three possible explanations.

10. Suppose we attempt to build a language model using a conditional probability table, where the probability of each word depends on all previous words:

$$P(\text{word}_n \mid \text{word}_1, \text{word}_2, \ldots, \text{word}_{n-1})$$

Assume the model is limited to a context window of 10,000 words and a vocabulary of 10,000 distinct words.

Explain the main problems and limitations that would arise with this design. Consider both computational and data requirements. Would this approach work in 20 years when computers are 1000000x faster and have 1000000x more memory? Why or why not?