Problem Set 1

Please submit a typed PDF addressing all problems below. This problem set contains 3 questions and is worth 25 points. **All responses MUST be in *your* own words.** Justification must be provided for **all** written answers. Statements made without any supporting explanation/justification will receive **no credit**. For mathematical derivations and plots, you may insert pictures of handwritten work if you find this easier. The required weekly readings and lecture slides should be helpful in completing the assignment. You can find these on our course website.

1. **Taxonomy of AI [2 points]:** For questions a and b, your answers should be ~2-4 sentences.

   (a) How is deep learning different from: (1) machine learning, and (2) artificial intelligence?

   (b) Give examples (not discussed in class/readings) of the following:

      i. An artificial intelligence algorithm which is not a machine learning algorithm.

      ii. A machine learning algorithm which is not a deep learning algorithm.

2. **Datasets and Tasks [3 points]:** For questions a, b, and c, your answers should be ~2-4 sentences.

   (a) Why are datasets used for deep learning often split into train and test sets?

   (b) Should there be any overlap of data between the training and test datasets?

   (c) How are classification and regression different?

3. **Artificial Neurons:**

(a) **Learning a Perceptron [5 points]**: Show the mathematical steps of learning a Perceptron model over one epoch, using the training data shown in Table 1. **This is NOT a programming exercise**. For full credit, you must include the mathematical steps used to derive the model parameters (bias and weights) and a table showing the model parameters after each training update. Recall that the perceptron model is of the form:

$$\hat{Y} = sign(\mathbf{w}^T \mathbf{x} + \mathbf{b}) \tag{1}$$

where $sign(\circ)$ returns the sign of the input, i.e. $sign(x) = x/|x|$ for $x \neq 0$. For training, use a learning rate of 0.25, and the following model parameter initializations:

$$\begin{aligned} \mathbf{w}^T &= [1, 1] \\ \mathbf{b} &= -1 \end{aligned} \tag{2}$$

Each training sample should be processed one at a time, in the order given by Table 1 (i.e. no random sampling).

(b) **Visualizing Model Testing [7 points]**: Provide a 2-D scatter plot of the testing dataset from Table 1. Plot X_1 along the "x-axis" and X_2 along the "y-axis". Use two distinct markers (e.g. $\times$ and $\circ$) to indicate the labels (Y). Then, using the result from part (a), overlay the model's decision boundary as a solid line. Recall that the decision boundary is the set of $\mathbf{x}$ values which satisfy:

$$\mathbf{w}^T \mathbf{x} + \mathbf{b} = 0 \tag{3}$$

Finally, report the model's predictions on the test data. You may hand-draw the plot and submit a picture in your PDF.

	Sample	X_1	X_2	Y
Training	1	1	1	-1
	2	-1	1	1
	3	-1	-1	1
	4	1	-1	-1
Test	1	0.5	2	1
	2	2	2	-1
	3	-0.5	0	1
	4	-0.5	-2	-1

Table 1: Training and test datasets.

(c) **Model Evaluation [4 points]**: Evaluate the prediction results for the final model on the test data by computing the confusion matrix, precision, recall, and accuracy.

(d) **Model Analysis [4 points]**: Discuss what insights are gained about the model's performance by examining each of the four evaluation results: confusion matrix, accuracy, precision, and recall.

(e) **Extra Credit [2.5 points]: Analysis of Learning Trajectory** Provide a 2-D scatter plot of the entire dataset (train and test combined), using two distinct markers to indicate the labels. Next, overlay the model's final/initial (i.e. after/before training the perceptron in part a) decision boundaries with solid/dotted lines, respectively. Then, conjecture a possible "optimal" solution for **w** and **b** which results in a model that correctly predicts all labels. Note that there are many possible such solutions. Overlay your chosen decision boundary as a dashed line on your plot. Finally, compare the model's initial and final decision boundaries to your chosen solution. Did the training process move the model's decision boundary closer to an "optimal" solution?

Collaboration versus Academic Misconduct: Collaboration with other students (or AI) is permitted, but the work you submit must be your own. Copying/plagiarizing work from another student (or AI) is not permitted and is considered academic misconduct. For more information about University of Colorado Boulder's Honor Code and academic misconduct, please visit the course syllabus.