Q1 Exam regulations

0 Points

BBM406 Honor Code

I promise that, for the BBM406 Midterm Exam

- All my exam work will be done entirely by myself, with no help from others;
- I will not communicate with anybody except the proctors during the exam;
- I will not consult any people or sources other than my printed/handwritten course notes, slides and the reference books listed in the course webpage;
- I will not provide any information about the exam's contents to other students until the exam deadline; and
- I will turn on my camera on Zoom session during the whole exam period.

Understanding this, I pledge my honor that I will not violate this Honor Code during the exam. I certify that all solutions will be entirely my own, that I will not consult people or sources other than those permitted, and that I will not share information with others during the exam.

Do NOT sign nor take this exam if you do not agree with this.

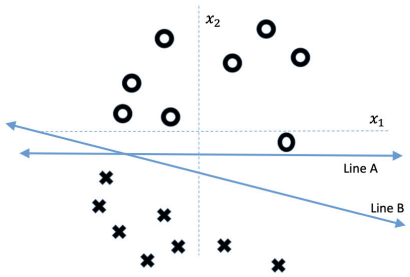
Signature (Specify your name and surname as your signature)

Mehmet Taha USTA MTUSTA

Q2 Logistic Regression

20 Points

Question 3. Logistic Regression [20 POINTS]
You have the following training set with each 2D sample point belonging to one of two classes ('x' denotes the positive class, and 'o' represents the negative class −1).



On this plot, line A and line B represent two different linear classifiers you obtained by using this training set.

(a) [5 POINTS] Assuming that line A is a horizontal line with an intercept of −1, and line B has a slope of −1/2 and an intercept of −2 on the vertical axis, please provide the parametric forms of these linear classifiers and some possible values of their model weights w.

(b) [5 POINTS] Suppose you use the following loss function for this problem:

$$J(w) = \sum_{i=1}^{n} \log\left(1 + \exp\left(-y_i x_i^T w\right)\right)$$

Please discuss the possible reasons on why the model parameters you found in part (a), either for line A or line B, won't actually be the same with the ones obtained by minimizing this loss function, *i.e.* $w^* = \arg\min \sum_{i=1}^{n} \log\left(1 + \exp\left(-y_i x_i^T w\right)\right)$. Why this is a problem, and how you can fix it?

(c) [5 POINTS] Now, suppose you include a term in the form of $\lambda(|w_1| + |w_2|)$ to the loss function given in part (c). Please discuss how increasing the value of λ affects the linear classifiers that are obtained by this combined loss function. Which one of the given lines, do you expect to have smaller loss value, line A or line B?

(d) [5 POINTS] Letting λ grow so large leads to the model parameters w_1 and w_2 to be very close to 0. In this case, what can you say about the final models obtained via minimization? What will be the final value of w_0 for the training data shown on the plot?

Please attach your solution to Question 3 (Logistic Regression) as a pdf file using the following link:

No files uploaded

Q3 Perceptron

20 Points

Question 4. Perceptron [20 POINTS]
The table below is a list of sample points in $\mathbb{R}^2$.

x_1	x_2	y
-4	2	+1
-2	1	+1
-2	-2	-1
2	2	-1
1	-2	-1

Suppose that you run the perceptron algorithm on these sample points considering the order given. Please answer the following questions accordingly.

(a) [10 POINTS] Suppose that the learning rate is $\varepsilon = 1$ and the initial weight vector is $w^{(0)}=(3, 2, 1)$, where the last component is the bias term. What is the equation of the separating line found by the algorithm, in terms of the features x_1 and x_2?

(b) [5 POINTS] Sometimes removing even a single point can change the decision boundary learned by the perceptron algorithm. For which, if any, point(s) in the given dataset would the learned decision boundary change if we removed it? Explain your answer.

(c) [5 POINTS] How would the result differ if the additional training point (3, 2) with label +1 is added to the training set?

Please attach your solution to Question 4 (Logistic Regression) as a pdf file using the following link:

No files uploaded

Q4 Convolutional Neural Networks

20 Points

Question 5. Convolutional Neural Networks [20 POINTS]

Consider a Convolutional Neural Network (CNN) that has an Input layer containing a 32×32× 1 image that is connected to a *convolution layer* having 10 5×5 filters with a stride of 1 and padding of 2. There is *no* activation function associated with the units in the convolution layer. This first *convolution layer* is connected to a *max pooling layer* using a 2×2 filter and a stride of 2 and padding of 0. Then, this *max pooling layer* is connected to a second *convolution layer* having again 10 5×5 filters with a stride of 1 and padding of 2, followed by another *max pooling layer* using a 2×2 filter and a stride of 2 and padding of. The *output layer* contains 10 units that each use an ReLU activation function and these units are fully-connected to the units in the pooling layer.

(a) [3 POINTS] Please specify the size of the activation volume at the first convolution layer? How many distinct *weights* must be learned for the connections to the first convolution layer?

(b) [3 POINTS] Please specify the size of the activation volume at the first pooling layer? How many distinct *weights* must be learned for the connections to the first pooling layer?

(c) [3 POINTS] Please specify the size of the activation volume at the second convolution layer? How many distinct *weights* must be learned for the connections to the second convolution layer?

(d) [3 POINTS] Please specify the size of the activation volume at the output layer? How many distinct *weights* must be learned for the connections to the output layer?

(e) [4 POINTS] Comment on whether CNNs can learn to recognize an object in an image no matter how the object is *translated* (i.e., shifted horizontally and/or vertically) even if the training set only includes that object in one position.

(f) [4 POINTS] Comment on whether CNNs can learn to recognize an object in an image no matter how the object is *rotated* (in the image plane) even if the training set only includes that object in one orientation.

Please attach your solution to Question 5 (Convolutional Neural Networks) as a pdf file using the following link:

No files uploaded

Midterm Exam - Part 2

GRADED

STUDENT

MEHMET TAHA USTA

TOTAL POINTS

0 / 60 pts

QUESTION 1

Exam regulations **0** / 0 pts

QUESTION 2

Logistic Regression **0** / 20 pts

QUESTION 3

Perceptron **0** / 20 pts

QUESTION 4

Convolutional Neural Networks **0** / 20 pts