Machine Learning, EX2

TO BE DONE INDIVIDUALLY. 4%.

--------------------------------------

Due Sun 10/22 at 11:59pm on Canvas.

Only .txt or .pdf are accepted. LaTeXing is recommended but not required.

This EX reviews HW1, and prepares for HW2.

------------------------------------------

Part I: Reflections from HW1

----------------------------

Although HW1 was a group coding exercises, we want to make sure each student

understands the entire code, and is able to hack it.

0. List your team members for HW1.

1. State your individual contributions to your group's HW1 submission. (Starting from HW2, your report will need to include author contributions).

2. Create your HW2 Group on Canvas.

3. Some groups included features that are not observed on the training data but might appear on dev/test, such as age=0, age=1, ..., age=100, etc., while other groups simply used observed features from the training set. Which option is better, or is there any difference?

4. You have observed that machine learning tends to exaggerate the existing bias in data. For example, there are about 24% positive examples in the dev set, but your models most likely predicted only around 20% positive on that set.

(a) is this due to overfitting or underfitting?

(b) is the exaggeration more severe on training set or on dev set? why?

(c) if we don't observe any feature, is it better to predict 100% or 75% negative? why?

(d) what other biases can you find that your best HW1 model exaggerates on the dev set? e.g., what about people with a Doctoral degree?

5. For your best model,

(a) what are the top five most positive/negative examples on the dev set?

(b) did your model predict them correctly?

(c) for each column (e.g., age), what is the most positive/negative feature?

(d) list five incorrectly predicted examples on the dev set. any observations?

Part II: SVM Theory

-------------------

1. State the two equivalent formulations for SVMs:

(a) maximize geometric margin

(b) minimize weight norm

2. What will happen

(a) if we replace the minimum functional margin of 1 by 10

(b) if we require the geometric margin to be at least 1

3. How many support vectors can there be for SVMs in d dimensions? (start with d=1,2,3)

4. True or False?

a) if an example has functional margin of 1, it must be in the support vector set.

b) if an example is in the support vector set, it must have a functional margin of 1.

5. List two reasons why the convex hull approach is not used to solve SVMs in practice.

6. Draw an example update of minibatch MIRA where the batch size is 2, i.e., you consider two examples at a time, and after the update, your new model should achieve a functional margin of at least 1 on both examples. Assume both examples are incorrectly predicted by the current model, and one example is positive while the other is negative.

(a) case one: one constraint is active.

(b) case two: both constraints are active.

(This problem helps you understand that MIRA update involves a special case of QP).

7. For a separable dataset, how many support vectors can there be for perceptron?

8. Why convex optimization is (much) easier than general optimization?

9. Are all quadratic programs instances of convex optimization?

Are all linear programs instances of convex optimization?

Why SVM is convex optimization?

10. Use Lagrangian multipliers to solve: min x^2, where 1<=x<=2.

Which constraints are active?

11. Explain the "complementary slackness" condition in KKT in your own words.

12. (optional) Why the alphas in Lagrange multipliers for inequality constraints must be non-negative?

13. (optional) Why minimizing the original constrained optimization objective becomes maximizing the new objective on alphas? (the optimal answer is achieved on a saddle point)

---------------------------------------

DEBRIEF SECTION (required)

1. Did you work alone, or did you discuss with other students?

If the latter please write down their names.

Note: in general, only high-level discussions are allowed.

2. How many hours did you spend on this assignment?

3. Would you rate it as easy, moderate, or difficult?

4. Are the lectures too fast, too slow, or just in the right pace?

5. Any other comments?

--------------- THE END OF EX2 ---------------------------------