

## HW03: Linear classification

Hand in at: <http://www.cs.utah.edu/~hal/handin.pl?course=cs5350>. Remember that only PDF submissions are accepted. We encourage using L<sup>A</sup>T<sub>E</sub>X to produce your writeups. See `hw01.tex` for an example of how to do so. You can make a `.pdf` out of the `.tex` by running “`pdflatex hw03.tex`”.

## 1 PRML Exercises

1. 3.4 (6350 only)
2. 4.7
3. \*4.14

## 2 Additional Exercises

1. The standard perceptron update is  $\mathbf{w} \leftarrow \mathbf{w} + y_n \mathbf{x}_n$  when a mistake is made. Suppose we were to add a learning rate  $\eta > 0$  so that the update became  $\mathbf{w} \leftarrow \mathbf{w} + \eta y_n \mathbf{x}_n$ . Does this change the final learned classifier? If so, how? If not, why not?
2. \*Suppose that instead of regularizing with an  $\ell_2$  norm, we were to regularize with an  $\ell_1$  norm. As we discussed in class, we can no longer use gradient descent, because the objective is not longer differentiable everywhere. Suggest a solution (you don't have to work through math: just tell me how you would approach this problem).
3. (6350 only) work through the math.