

## HW10: Ensemble methods, clustering

Hand in via moodle at: <https://moodle.umass.edu/course/view.php?id=20836>. Remember that only PDF submissions are accepted. We encourage using L<sup>A</sup>T<sub>E</sub>X to produce your writeups. See `hw00.tex` for an example of how to do so. You can make a `.pdf` out of the `.tex` by running “`pdflatex hw00.tex`”. You’ll need `mydefs.sty` and `notes.sty` which can be downloaded from the course page.

1. For bagging, why is it important to do bootstrap resample (sampling with replacement) rather than sampling without replacement?
2. Suppose you are training a binary classifier by boosting decision stumps where all features are binary. In other words the weak learners are decision trees of depth one, i.e., consider only one variable. Show the resulting classifier after boosting has the same form as a linear classifier, i.e.  $f(\mathbf{x}) = \text{sign}(\mathbf{w}^T \mathbf{x})$  for some weight  $\mathbf{w}$ .
3. What happens to the objective function for K-means as K increases? Why?