Extra Credit Homework

Issue Date: April 8, 2019
Due Date: April 15, 2019

We're issuing this optimal assignment for those who may want to improve your total homework grade. This homework is completely optional meaning you can still get 100% in the course if you choose not to do this extra work. However, doing this assignment can improve your homework grade. Note that this homework is **not** meant to completely replaced any of your homeworks either. Finally, your score on this extra assignment does **not** contribute in any way to your project score.

If you choose to do the assignment, your answers must be submitted electronically via CAMU. You must provide a single **PDF** containing all your answers. The PDF file can be generated using editing software (e.g., Latex) or can be a scanned copy of your hand-written answers. If you opt for scanned copies, please make sure to use a scanner, not your phone camera. Familiarize yourself with the process of scanning documents a few days before the due date. Late submissions due to scanning problems will be treated and penalized as any other form of late submission, as per the rules stated in the Syllabus. Also, make sure your scanned submission is readable. If we cannot read it clearly, you will receive 0 points for that problem.

1. [**25 points**] We've covered both linear regression and logistic regression in class. Compare and contrast these two learning algorithms. Remember to be as clear, precise and concise as possible.

2. [**50 points**] Dr. Korsah taught you reinforcement learning extensively just before the mid-semester break. Explain in *detail*

   (a) what reinforcement learning is

   (b) the differences between reinforcement learning, supervised learning and unsupervised learning

   (c) the similarities and differences between classical reinforcement learning and deep reinforcement learning

   (d) why supervised learning is still the predominant form of machine learning when reinforcement learning has the potential to yield better results on many tasks.

3. [**25 points**] In the k-means lecture Dr. Korsah delivered just after the mid-semester break, you were introduced to another optimization method: coordinate descent. Briefly explain the difference(s) between gradient descent and coordinate descent if any. Why does k-means use coordinate descent instead of gradient descent?