

The Midterm



CSC411/2515: Machine Learning and Data Mining, Winter 2018

Michael Guerzhoy and Lisa Zhang

Logistics

- Main sitting: Friday March 2nd, 6pm-8pm
- If you have a conflict, please email us ASAP
- Link on Piazza to register for a 7pm-9pm sitting – register by Feb. 16!
- You should write in pen if you want to potentially contest your grade on the midterm
 - We will not accept remark requests for midterms written in pencil

Coverage

- Everything up to and including the EM tutorial
 - We will not ask about PCA and decision trees on the midterm
- Lecture and tutorial content
- Projects 1 and 2
 - Techniques rather than remembering detail
- Study guide problems
 - At least one midterm question will be very closely related to a study guide problem
- Problem solving in the context of ML
 - Thinking the lectures through should help

Study Guide

- It looks daunting, but it should *save* you time
 - Problems that are asking you what a slide or two in the lecture say
 - Problems that are asking you to apply lecture content in a straightforward way
 - A few problems that require problem-solving
- There are a lot fewer problems than there people in the class
 - If everyone contributes the equivalent of 2 solutions to the Google Doc, we'll be in good shape

Sample midterms

- Winter 2017:
 - <http://courses.skule.ca/course/CSC411H1>
- Fall 2017 practice midterm:
 - http://www.cs.toronto.edu/~jlucas/teaching/csc411/resources/example_midterm.pdf
 - http://www.cs.toronto.edu/~jlucas/teaching/csc411/resources/example_solutions.pdf
 - Skip 1a, 1c, 3a, 3b. 4 is doable but quite difficult, especially without matrix calculus.

Formulas

- No cheat sheets
- Any formula that we did not derive or explain completely will be provided to you. For example, if you need the following, they will be provided:
 - The closed form solution for linear regression, univariate/multivariate Gaussian density, tanh, logistic regression model, the definition of variance, Xavier initialization
- Formulas which we did derive may not be provided, depending on the problem
 - We derived the log-loss
- Formulas that we explained or that are very straightforward will not be provided
 - ReLU, sigmoid, mean square loss

Study advice

- Our goal is for you to understand everything in the lectures, tutorials, and projects, and be able to apply it
 - Study guide questions are intended to help you go through the process of making sure that you understood everything
 - Make up your own study guide questions
- Go to study guide sessions/contribute to the Google Doc