CSCI 5521: Machine Learning Fundamentals (Spring 2025)

Quiz 1 (Thursday, Feb 13)

Due on Gradescope at 02:00 PM, Friday, Feb 14

Instructions:

- This quiz has 3 questions, 30 points, on 1 page.

- Please **write your name & ID on your submission pages**.

1. (**6 points**) Self-driving cars rely on machine learning models to interpret weather conditions and make real-time driving decisions. One important task is to use camera and sensor data onboard to classify whether it is snowing outside or not, which helps the vehicle adjust its speed and braking behavior for safety. Please model this problem as a two-class classification task and answer the following questions.

   (a) Name two relevant features to this two-class classification task (i.e., whether it is snowing outside or not). Any reasonable features are acceptable.

   (b) What are the labels in this two-class classification task?

2. (**10 points**) What could we do to reduce overfitting in a polynomial regression model? Select all the option(s) that apply.

   (a) Remove all training data that is incorrectly classified by the model during the first iteration.

   (b) Decrease polynomial degree of the model.

   (c) Increase polynomial degree of the model.

   (d) Collect 20% more new training data.

   (e) Collect 30% more new test data.

3. (**14 points**) The conditional probability density functions of two classes C_1 and C_2 are shown in the figure below, with $P(x|C_1) \sim \mathcal{N}(0.5, 1.3)$ and $P(x|C_2) \sim \mathcal{N}(1.1, 0.7)$.

   (a) Assuming the priors are equal, predict which class (C_1 or C_2) the data point $x = 1.7$ (illustrated with the green dot) belongs to. Briefly explain why.

   (b) What if the priors are $P(C_1) = 0.6$ and $P(C_2) = 0.4$ respectively?