* This is the open-notes exam. Please do not use the Internet and work individually.
* You are welcome to use any packages and libraries.
* Compress all files into a single zip file and submit. When you submit, name the file as “yourlastname\_CS559\_finalexam.zip”.
* Submit by 5/9th Monday 11:59 PM.

Q1: [20 pts] KNN-Clustering: Use Q1.csv.

Suppose we are building a model that clusters using KNN.

1. [10 pts] Build a KNN model that iteratively clusters by the new classes.
2. [5 pts] Visualize the number of cluster ID population changes in each cluster.
3. [5 pts] Use KMeans to cluster and estimate the accuracy of the model from a).

Q2: [30 pts] Random Walk Probability.

Considering a robot that travels from the origin (0,0) to an exact target (2,2). Suppose the robot recognizes the direction by the signal between 0 and 1 as follows:

Assume the robot will crash if its either coordinate exceeds 2.

1. [7 pts] Numerically calculate the probability of arriving at the target.
2. [12 pts] Build a simulator that projectiles the position of the robot for 20 sequential signals.
3. [8 pts] Generate a data set and demonstrate the success of probability for 200 simulations with an increment of 50 trials in each simulation.
4. [5 pts] Using the generated data set, demonstrate the one success case graphically.
5. [8 pts] Use the generated data from c). Build a simple neural network approach success classifier after 4th movement. Hint: make a binary test (either closer to the target or farther away from the target) in each layer after the 4th movement. Do not implement the full neural network. Instead, do similar to the “car accident” example shown in the lecture note.
6. [5 pts] To improve the success, we are going to use the Gini index to fix the signals at three movements from the hidden layers made from d). Report the i-th number of movement and signal values. Did the simulation improve after manual control on three i-th movements?

Q3: [25 pts] Gradient Descent.

Consider a function damping oscillation function

where and are constant, is the noise in , and the domain of function is .

1. [5 pts] Starting with constant set in the order of and , show that the function converges using gradient descent. Use a dataset Q3.
2. [5 pts] Using the found constant from a) and analytically solve for when . Hint:
3. [5 pts] Using the Linear Regression, find the rate of decay function .
4. [5 pts] Using the Gaussian Process, forecast the next three local maxima. Report the answer as . Hint: once , changes its sign.

Q4: [10 pts] Softmax derivation.

Show the derivative of the error function

with respect to the activation for output units having a softmax activation function satisfies

where and the network outputs are .

Q5: [15 pts] Graphical Models

Show that the marginal distribution for the variables in a factor in a tree-structured factor graph, after running the sum-product message passing algorithm, can be written as the product of the message arriving at the factor node along all its links, times the local factor in the form