ID 5030 Quiz -2

1. Deadline: Sunday at 1 pm (afternoon). No late submissions will be allowed under any circumstances.

2. Sign the honour code declaration and upload it too.

3. You are allowed to refer to books, websites, open source codes etc but mention all references clearly.

4. You are not allowed to discuss the quiz with anyone else. This will be cause for a U grade in the course.

5. **Make any assumptions needed and state them explicitly if you find the question incomplete.**

6. All derivations should have every step clearly written

7. Make the answers to descriptive questions short and precise.

8. Questions will be dynamically weighted. That is, answers that are solved correctly by the least number of students will have the greatest weight.

---

1. Download the Old Faithful data from the following link
   http://www.stat.cmu.edu/~larry/all-of-statistics/=data/faithful.dat.

   a. What are the various ways of clustering this data?

   b. Write a simple program to cluster the data. If you wish, you may use any available open source code as long as you provide a clear reference.

   c. Using the clustered data from the previous part, train a Naïve Bayes classifier.

   d. For the extracted clusters from part (b) - fit maximum likelihood Gaussians to the individual clusters. Derive any expressions that you use

   e. What would you need to do in order to provide MAP and Bayesian estimates? Derive the expression for the same. You need not give any final values, just the derivation and final expressions are sufficient.

2. Consider the "Inception" module as discussed in https://arxiv.org/pdf/1409.4842.pdf.

   a. What is/are the advantage/s of using an Inception module?

   b. Consider Figure 2 in the paper. Which of the two Inception models shown there are better and why?

   c. Consider the AlexNet architecture discussed in class. Suggest a simple way to incorporate the Inception module idea within AlexNet.

   d. List a couple of applications of an encoder-decoder type architecture. How will you incorporate Inception module in the encoder and decoder sides of an encoder-decoder type network?

3. An important open problem in Deep Learning is the interpretability of deep neural networks. That is, after training, it is often difficult to interpret for a human why the network gave the output that it did.

   In this context, let us say that a deep CNN is used to predict lung tumours from X-ray radiography images. However, the radiologists are sceptical of a black box prediction from a CNN and would like understand how the network arrived at the output. Suggest a method by which Binary decision trees could be combined with CNNs to address the radiologist's concerns of being able to interpret the output.

   [P.S -- https://bit.ly/2otrSjZ has a detailed popular article on the interpretable deep learning problem. The article is just for your interest and is not necessary to read for attempting this question]

4. During a product survey involving two competing soft drink brands A and B, N persons were asked for their preferred drink out of which n_A people indicated their preference for A. Derive expressions for the probability that a general person in the population prefers drink A using the ML, MAP and Bayes estimates. Assuming a prior distribution in the population to be a beta distribution with both shape parameters to be 5.

   [Do not be intimidated. We have not covered beta distributions in class but the expressions are readily available here https://bit.ly/2JGcito]

5. For the soft margin linear classifier (that is, a linear classifier which accounts for possible outliers), write down the corresponding primal problem. Derive, listing all steps, the expression for the dual problem and the corresponding KKT conditions.

6. We wish to write a numerical solver for a general, first order ODE of the form dy/dt = f(y). We are given the function f(y) and also the initial condition y(t=0) = y0. The usual method of solving this involves using a numerical integration technique such as the 4th order Runge-Kutta method (RK4). [See https://bit.ly/2GURhcB]

   Design a vanilla RNN architecture to perform a RK4 integration.