Midterm Exam for Machine Learning, 5/8/2020

Please show your work. That is, you need to provide the intermediate steps toward the answers. I cannot accept a reason like "the answer is directly obtained from a software package, so I don't know how to calculate it myself." It is OK to write a simple program with online resources such as Octave (an open source version of MATLAB) to save calculation time. But, you still need to hand-calculate at least one value to show how to use the equations.

1. Based on the table below to construct a Naïve Bayesian classifier. If a particular day is sunny, high temperature, low humidity, and no wind, what is the decision? (15 points)

Outlook	Temperature	Humidity	Windy	Decision
Sunny	Hi	Hi	No	No play
Sunny	Hi	Hi	Yes	No play
Overcast	Hi	Lo	No	Play
Overcast	Lo	Lo	Yes	Play
Rain	Lo	Hi	No	Play
Rain	Lo	Hi	Yes	No play

2. Repeat problem 1 but make a decision based on ID3 decision tree. (15 points)
3. Repeat problem 1 with a 3-NN classifier. For each attribute, if two values are the same, their distance is zero. Otherwise, their distance is one. For example, the distance between Yes and No is one, but the distance between Sunny and Sunny is zero. If two distances are equal, the one with same outlook takes precedence. (10 points)
4. Suppose that we have a dataset containing samples $[x_1 \quad x_2]^T$ from jointly Gaussian random variables with $\mu = \begin{bmatrix} 2 \\ 4 \end{bmatrix}$ and $\Sigma = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$. Is it possible to reduce these two attributes into one by using PCA? Use $\text{PoV}(1) > 0.9$ as the criteria to answer this question. Can answer this question with only intuition (i.e., without computing eigenvalues)? (10 points)
5. Analytically solve the following problem: Minimize $f(x, y) = x^2 + y^2$ subject to $y = x^2 + 1$. (15 points)
6. The following one-dimensional data points are known from two clusters: [0.9, 0.7, 1.2, 2.4, 1.8]. Use hand calculation to perform one epoch of k-means algorithm. The initial conditions are $\mu_1 = 1$ and $\mu_2 = 2$. (15 points)

7. For the neural network given below, the activation function from q_1 to h_1 and q_2 to h_2 is sigmoid, the output nodes have linear activation function (i.e., $z_i = y_i$), and the cost function is $J = \sum_{i=1}^{2}(y_i - d_i)^2$. Let w_1 to w_8 be 1.0, $d_1 = 0.9$, $d_2 = 0.1$, and $\eta = 0.1$.

   i. Find y_1 and y_2 (forward computation) if $x_1 = x_2 = 0.5$. (5 points)

   ii. Find the value of $\Delta w_1 = \eta \frac{\partial J}{\partial w_1}$ using backprop if $x_1 = x_2 = 0.5$ (10 points)