**11482 Pattern Recognition and Machine Learning  
11512 Pattern Recognition and Machine Learning PG**

**Tutorial 5 – Week 6 Discussion Questions (Lab is attached separately)**

1. Which type of machine learning system should you use to make a robot learn how to walk? Briefly explain.
2. Suppose you have trained a polynomial model y(x) = w0+w1x+...+wMxM via least-squares regression, and you find that it has a low training error, but a high testing error. Which of the following is likely to reduce the testing error? (a) Increasing the number of training data points and re-training. (b) Decreasing M and re-training. (c) Dividing each of the learnt weights by 2. (d) Increasing M and re-training.
3. When training a multi-variable regression model, which of these would be a reasonable way to gauge the importance of different features? (a) Weights with higher values will correspond to more important features. (b) Normalise each feature to have mean 0 and variance 1; then weights with higher values will correspond to more important features. (c) Normalise each feature to have mean 0 and variance 1; then weights with higher absolute values will correspond to more important features. (d) The greater the increase in cross-validation error on eliminating a feature, the more important it is.
4. After running 3 fold validation on your data, you got the following cross validation results.

Table

Description automatically generated

What is the cross validation accuracy for the result? When should you use training accuracy and how is test accuracy used?

1. Given a dataset with two classes A and B, where 80% of the samples belong to class A and 20% belong to class B. If you have a feature vector x, calculate the posterior probabilities P(A|x) and P(B|x) using Bayes' theorem.
2. Why is Naïve Bayesian called ‘naïve’?
3. Briefly explain the ‘zero probability/frequency’ problem in Naïve Bayesian classifiers? How can this be resolved?
4. After your medical checkup, the doctor has bad news and good news. The bad news is that you tested positive for COVID, and that the test is 99% accurate (i.e., the probability of testing positive given that you have COVID is 0.99, as it is the probability of testing negative given that you do not have the disease). The good news is that this is a rare disease, striking only one in 10,000 people. Why is it good news that the disease is rare? What are the chances that you actually have COVID?