General Instructions: Write or type your answers neatly and remember to show all relevant work. All questions are worth 10 points. Each answer should be a separate pdf, and you can turn in the pdfs on canvas in the appropriate assignment. Some questions may be very challenging; significant partial credit is available for reasonable attempts at solutions. Since each question is worth the same number of points, do not waste too much time on any one. Ask me or the TAs for help if stuck.

Some of the questions require you to write short programs to simulate things. You can use any language/software to do this, and you do not need to turn in your code.

Upload your answers to Canvas as a pdf file by 11:59pm on the due date specified after the question. You will receive a 10% bonus for a solution turned in a week or more in advance of the due date. You can use one late day each week (up to Saturday 11:59pm) with a penalty of 20%. Submissions after Saturday 11:59pm for any week will not be graded.

Each group must do their own work. Only one submission is needed from each group. Do not use any source other than the lecture notes, textbook(s) and readings on the class website to answer these questions. Only those who contributed equally to a submission should have their names and Case IDs on the submission. Those not listed as contributing will not receive points.

33. Consider a modified SVM formulation derived using the plus-plane at $\mathbf{wx}+b=c_1$ and the minus-plane at $\mathbf{wx}+b=c_2$, $c_1>0$, $c_2<0$. Explain the relationship between the decision surfaces obtained when (i) $|c_1|>|c_2|$, (ii) $|c_1|<|c_2|$, (iii) $|c_1|=|c_2|$. When would we prefer one over the other?

34. Explain why the margin of classification in an SVM $(\mathbf{w},b)$ is independent of b.

35. Two classifiers A and B are evaluated on samples of size n and found to have error rates e_A and e_B such that $e_A-e_B=0.1$. If the true error rates of A and B are indeed different, how large does n have to be to guarantee we can establish the difference at a 95% confidence level?

36. A revolutionary new classifier, the Deep Bayesian Logistic Neural Tree Kernel Network, has been invented. Professors Bayesian Bob and Neural Nan have independently evaluated such a classifier on two datasets from a learning problem, obtaining 95% confidence intervals of (x_B, y_B) and (x_N, y_N) respectively. Over dinner, they share their findings with each other. Unfortunately, they are overheard by Professor Scoop, who wants to publish the result without doing the experiment. Find an expression for the best 95% confidence interval that Scoop could derive from Bob and Nan's findings. You can assume the 95% CI multiplier for the Gaussian distribution to be 2 for convenience.