

Assignment 3

75 marks

Remember to comment your code well, in order to get marks reserved for comments

Q1. We saw the perceptron algorithm in Lecture 15. I want you to implement this algorithm, and use it to classify banknotes as forged or authentic in this [dataset](#). In particular,

- (a) Code up the perceptron algorithm described on slide 7 of Lecture 15 using the same notation as in the slides. [10 points]
- (b) Write functions to make predictions using the algorithm for the banknotes dataset. Preprocess the dataset to handle missing and anomalous data. [10 points]
- (c) Train the algorithm on the dataset using cross-validation and report cross-validated test set error [10 points]
- (d) Ensure you use a held out validation set and report F1 score on the held out set for your best model [5 points]

Q2. In Lecture 14, we saw how we can use MCMC sampling to approximate Bayesian posteriors when the prior and likelihood distributions are not conjugate. Let's consider a simple demonstration of MCMC sampling in a setting where conjugacy is actually possible – normal likelihoods with a known population variance, for which the prior is another normal distribution.

- (a) Write a function to calculate the Bayesian posterior probability given 50 new data samples drawn from a normal distribution with mean 10 and SD 5, assuming a normal prior with mean 25 and s.d. 5. Plot the pdfs of the prior, the likelihood and the posterior distributions. Explain how you derive the likelihood from the data. [15 points]
- (b) Implement the Metropolis algorithm from the lecture slides to estimate the posterior distribution given the same prior and data and show that it converges to the analytic posterior by plotting a histogram of samples from the distribution alongside the analytic posterior distribution. Assume whatever SD (width) you want for the proposal distribution. [15 points]
- (c) How does the speed of convergence of the sampling depend on the proposal width? Is there an optimal proposal width that would work best? Demonstrate the consequences of using sub-optimal proposal width and terminating sampling too soon. [10 points]