AI1001: Introduction to Modern AI
Final Examination
28 August 2019, 12:00 – 13:30

Attempt as many questions as you can. All questions carry equal weight.

1. Suppose A and B are Boolean variables, and that $\Pr\{A \vee B\} = 0.9$, while $\Pr\{A \wedge B\} = 0.6$. Find the maximum and minimum values of $\Pr\{B\}$ that are consistent with this information.

2. Recall the linear programming approach to computing the maximum and minimum probabilities of various Boolean formulas, when the data consists of the probabilities of other Boolean formulas (as in Problem 1 above). Explain why this approach is limited to only a small number of Boolean variables, and does not scale to a large number of Boolean variables.

3. Explain what a support vector machine is in your own words.

4. Explain the rationale behind a higher-order support vector machine.

5. Explain what logistic regression is in your own words.

6. Suppose you are given a set of labelled training samples $\{(\mathbf{x}_1, y_1), \ldots, (\mathbf{x}_m, y_m)\}$, where each label y_i is bipolar (that is, equals ± 1). Suppose you wish to use logistic regression, which requires binary and not bipolar labels. How would you handle this situation?

7. What is the difference between a feedforward neural network (FFNN) and a recurrent neural network (RNN)? What can a RNN do that a FFNN cannot do?

8. Suppose a Markov process with five states has the following state transition matrix:

$$\begin{bmatrix} 0.1 & 0 & 0.3 & 0.4 & 0.2 \\ 0.1 & 0.1 & 0.7 & 0 & 0.1 \\ 0.2 & 0.3 & 0.4 & 0.1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}.$$

   - Suppose that at time $t = 0$, the Markov process is in state 2. What is the probability distribution of the state at time $t = 1$?
   - Are there any absorbing states, and if so, which are they?
   - What do you think the long-term behaviour of the Markov process would look like? Justify your answer.

9. A Markov process has the following state transition matrix:

$$A = \begin{bmatrix} 0.4 & 0.5 & 0.1 \\ 0.3 & 0.4 & 0.3 \\ 0 & 0 & 1 \end{bmatrix}.$$

   For an absorbing state, the "hitting time" vector is given by $(I - \bar{A})^{-1}\mathbf{1}$, where $\bar{A}$ is the state transition matrix A without the row and column corresponding to the absorbing state, and $\mathbf{1}$ is the vector of all ones. Compute the average number of time steps needed to reach the absorbing state from the other states.

10. A Markov process has the following state transition matrix:

$$A = \begin{bmatrix} 0.4 & 0.5 & 0.1 & 0 \\ 0.3 & 0.4 & 0 & 0.3 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

- There are two absorbing states; what are they?
- To find the probability of winding up in an absorbing state, define $\bar{A}$ to be the submatrix of A without the absorbing states, and $\mathbf{y}$ to be the column of A corresponding to the absorbing state. Then the probability of hitting the absorbing state is $(I - \bar{A})^{-1}\mathbf{y}$. With this refresher, compute the probability of ending up in each of the absorbing states starting from the other states.