AI (Fall 2019) – Assignment 4
Machine learning

Due: 11:59pm, Wednesday, Dec. 25, 2019

1. Consider the following data. The DECISION-TREE-LEARNING algorithm will first select the attribute Length to split on. Finish building the decision tree, and show the computations.

Example	Author	Thread	Length	Where Read	User Action
e1	known	new	long	home	skips
e2	unknown	new	short	work	reads
e3	unknown	follow Up	long	work	skips
e4	known	follow Up	long	home	skips
e5	known	new	short	home	reads
e6	known	follow Up	long	work	skips
e7	unknown	follow Up	short	work	skips
e8	unknown	new	short	work	reads
e9	known	follow Up	long	home	skips
e10	known	new	long	work	skips
e11	unknown	follow Up	short	home	skips
e12	known	new	long	work	skips
e13	known	follow Up	short	home	reads
e14	known	new	short	work	reads
e15	known	new	short	home	reads
e16	known	follow Up	short	work	reads
e17	known	new	short	home	reads
e18	unknown	new	short	work	reads

2. Consider the candy example from the lecture. Assume that the prior distribution over $h_1, \ldots, h_5$ is given by $\langle 0.1, 0.2, 0.4, 0.2, 0.1 \rangle$. Suppose that the first 5 candies taste lime, cherry, cherry, lime, and lime. Make predictions for the 6th candy using Bayesian, MAP and ML learning, respectively. Show the computations done to make the predictions.

3. Consider the Boolean function E = (A XOR B) AND (C XOR D). Construct its truth table, and then remove the line for the input $A = 1, B = 1, C = 1, D = 1$. Use Naive Bayes classification to make prediction for this input. Show the computations.

4. Construct a neural network that computes the XOR function of two inputs.

5. Consider the neural net on Page 32 of the course slides for neural nets.

   - Suppose we use the sigmoid function as the activate function. Compute $\partial Loss_{o_1}/\partial w_1$.
   - Suppose we use the tanh function as the activate function. Compute $\partial Loss_{o_2}/\partial w_4$. Note that $tanh(x) = (e^x - e^{-x})/(e^x + e^{-x}) = 2g(2x) - 1$, and $tanh'(x) = 1 - tanh^2(x)$.