CS 1571 Last “Homework” (not to turn in)

Problems from the text:

16.8

18.6, 18.11

Solutions follow

16.8 The expected monetary value of the lottery L is

$$EMV(L) = \frac{1}{50} \times \$10 + \frac{1}{2000000} \times \$1000000 = \$0.70$$

Although $\$0.70 < \$1$, it is not *necessarily* irrational to buy the ticket. First we will consider just the utilities of the monetary outcomes, ignoring the utility of actually playing the lottery game. Using $U(S_{k+n})$ to represent the utility to the agent of having n dollars more than the current state, and assuming that utility is linear for small values of money (i.e., $U(S_{k+n}) \approx n(U(S_{k+1}) - U(S_k))$ for $-10 \leq n \leq 10$), the utility of the lottery is:

$$\begin{aligned} U(L) &= \frac{1}{50}U(S_{k+10}) + \frac{1}{2,000,000}U(S_{k+1,000,000}) \\ &\approx \frac{1}{5}U(S_{k+1}) + \frac{1}{2,000,000}U(S_{k+1,000,000}) \end{aligned}$$

This is more than $U(S_{k+1})$ when $U(S_{k+1,000,000}) > 1,600,000U(\$1)$. Thus, for a purchase to be rational (when only money is considered), the agent must be quite risk-seeking. This would be unusual for low-income individuals, for whom the price of a ticket is non-trivial. It is possible that some buyers do not internalize the magnitude of the very low probability of winning—to imagine an event is to assign it a "non-trivial" probability, in effect. Apparently, these buyers are better at internalizing the large magnitude of the prize. Such buyers are clearly acting irrationally.

Some people may feel their current situation is intolerable, that is, $U(S_k) \approx U(S_{k\pm1}) \approx u_\perp$. Therefore the situation of having one dollar more or less would be equally intolerable, and it would be rational to gamble on a high payoff, even if one that has low probability.

Gamblers also derive pleasure from the excitement of the lottery and the temporary possession of at least a non-zero chance of wealth. So we should add to the utility of playing the lottery the term t to represent the thrill of participation. Seen this way, the lottery is just another form of entertainment, and buying a lottery ticket is no more irrational than buying a movie ticket. Either way, you pay your money, you get a small thrill t, and (most likely) you walk away empty-handed. (Note that it could be argued that doing this kind of decision-theoretic computation decreases the value of t. It is not clear if this is a good thing or a bad thing.)

18.6 In order to find the best test to split the data, the algorithm finds the attribute that leads to the largest information gain, then recurses on the resulting pieces of the data set.

The first test selected is on A1. One side of the split is a pure node, meaning there is only one label in the node. The other set of instances will then be split on A2 , and every leaf in the resulting tree will be a pure node.

18.11 The majority classifier simply selects the most common label among the training instances. If a positive example is left out, there will be 49 positive training instances and 50 negative ones so the classifier will predict negative. When a negative instance is left out, there will be more positive training instances, so the classifier will predict positive. Therefore, the classifier will never make a correct prediction using leave-one-out cross-validation.