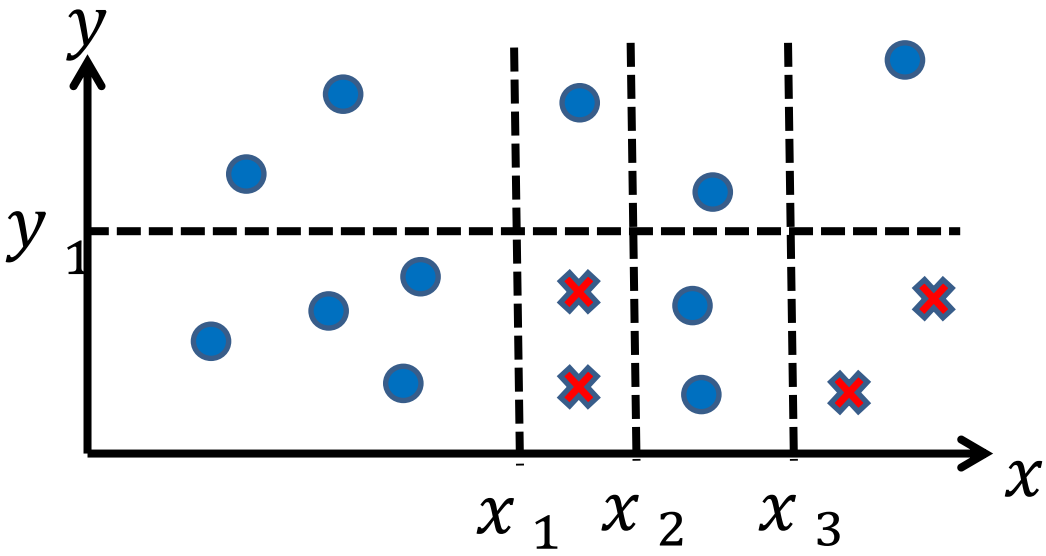
HW #4 Due: 4/16/2021

1. Follow the model in the HMM file (pp. 13) to find the most probable sequence of states Q (decoding problem). The observation sequence is "RRB" (R: red, B: blue). Assume $\pi = \left[\frac{1}{3} \quad \frac{1}{3} \quad \frac{1}{3}\right]$ and all elements in A are $\frac{1}{3}$.

2. Plot a decision tree for the following data points. If you carefully design your tree, you will just need to use one ">" or "<" in a vertex.



3. We have G = 0.048 for H_0 = 65 and G = 0.102 for H_0 = 80 on pp. 46 of the decision-tree PPT file. Confirm these G values are correct by hand calculation.
4. Write a program for multinomial HMM classification. The training and test sequences are given in HMM_data.npz and the sample program HMM_read.txt shows how to read the data.
5. Use the data below to construct a CART tree and plot the resultant tree using sklearn (or whatever tools you plan to use). You need to think a way to deal with the categorical data. If a particular day is sunny, high temperature, low humidity, and no wind, what is the decision based on your plotted tree?

Outlook	Temperature	Humidity	Windy	Decision
Sunny	Hi	Hi	No	No play
Sunny	Hi	Hi	Yes	No play
Overcast	Hi	Lo	No	Play
Overcast	Lo	Lo	Yes	Play
Rain	Lo	Hi	No	Play
Rain	Lo	Hi	Yes	No play