CS-661 AI Assignment 2

1. A linearly separable 2-class classification is one for which there exists a vector $\mathbf{v}$ and a scalar θ such that for every point $\mathbf{x}$ in the first class, $\mathbf{v} \cdot \mathbf{x} > \theta$ while for every point in the second class $\mathbf{v} \cdot \mathbf{x} < \theta$.
For example, for two inputs which are constrained to be $\{0, 1\}$ we can define two classes according to whether the **and** function on the two coordinates returns 0 or 1. Thus the first class contains the points $(0, 0)$, $(0, 1)$ and $(1, 0)$ since $0 \wedge 0 = 0 \wedge 1 = 1 \wedge 0 = 0$ while the second class contains only $(1, 1)$ since $1 \wedge 1 = 1$. This classification is linearly separable. Find the $\mathbf{v}$ and θ.
The **or** function similarly defines a linearly separable classification, with $(0, 0)$ in the first class, since $0 \vee 0 = 0$, while all the others in the second class $0 \vee 1 = 1 \vee 0 = 1 \vee 1 = 1$. What are $\mathbf{v}$ and θ now?
The function **xor** defined by $0 \otimes 0 = 1 \otimes 1 = 0$ and $0 \otimes 1 = 1 \otimes 0 = 1$ defines a non-linearly separable classification. Prove this. (This is essentially Minsky and Papert's 1969 argument against neural networks, which prevented funding of neural research for 15 years.)

2. A family of classifiers is said to *shatter* a set of points if no matter how we divide the points into two classes, the family contains a classifier which properly realizes this classification. The V.C. dimension, d^{VC}, of a family of classifiers is defined to be the largest number of points which can always be *shattered*. This dimension is a measure of the 'strength' of the classifier family.
Prove that for points on the line, the family of all finite intervals has a V.C. dimension of 2.
Show that for points in a plane, the family of all rectangles with sides parallel to the axes has $d^{VC} = 4$.