Question 2

And could be achieved by putting 0.5 for the weight and -0.7 as theta. Negative could be put as -1 and theta as 0.7.

For XOR, it’s ‘or’ – ‘and’\*2. So, the second layer weight for A and B is 0.5 and for the ‘and’ is -2 and theta is 0.

Question 4

Instead of using gini index or entropy, it selects the feature to split by minimizing the variance of the target variable. At the leaf node, it takes the mean of the target variable as the prediction for the sample.

Question 6

Decision tree should perform better than the nearest neighborhood.

KNN would suffer for the points lying on the or around the lines.

On the other hand, with two features and a theta, decision tree can describe the exact line on the plane.

Question 7

1. Circle should have 3 VC dimensions. Since 3 points confirms a circle, then if the 4th point is within the triangle, there’s no way to exclude the point
2. Sphere should have 4 VC dimensions. 4 points confirms a sphere, then if the 5th point is with the sphere, there’s no way to exclude the point