CES 417T - HW6

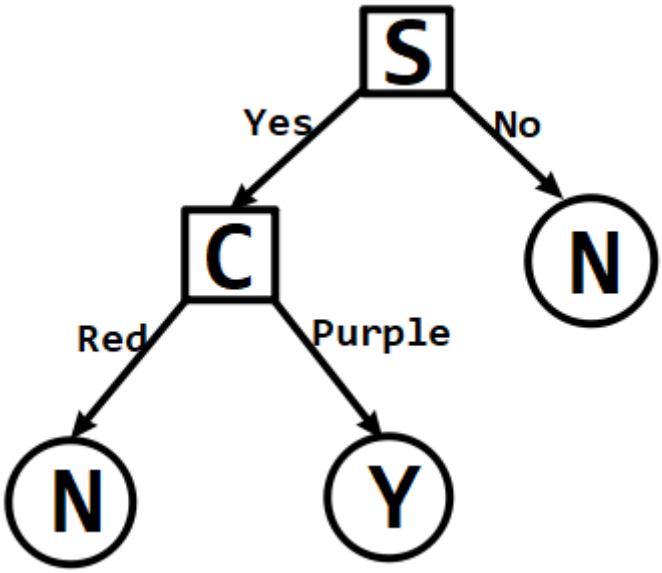
467261 - Yifu Wang

2018 - 12 - 01

1. Let denote C, S, T, P as Color, Stripes, Texture, Poisonous.

(a) $Entropy(P) = -\frac{3}{5}log_2\frac{3}{5} - \frac{2}{5}log_2\frac{2}{5} = 0.673$
$IG(C, P) = Entropy(P) - \frac{4}{5}(1) - \frac{1}{5}(0) = -0.127$
$IG(S, P) = Entropy(P) - \frac{2}{5}(0) - \frac{3}{5}(-\frac{1}{3}log_2\frac{1}{3} - \frac{2}{3}log_2\frac{2}{3}) = 0.291103$
$IG(T, P) = Entropy(P) - \frac{3}{5}(-\frac{1}{3}log_2\frac{1}{3} - \frac{2}{3}log_2\frac{2}{3}) - \frac{2}{5}(1) = -0.109$
So the root attribute of the tree will be Stripes.

(b) Decision tree

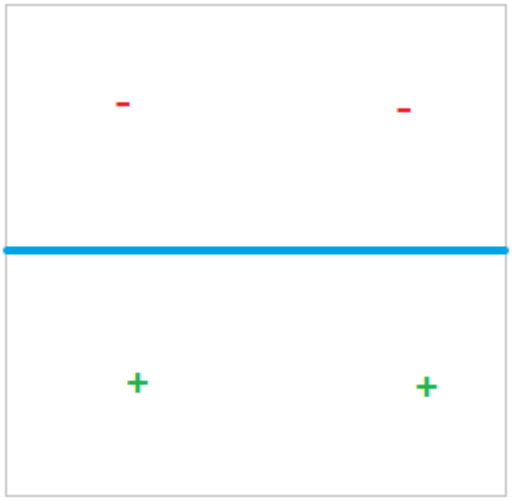


2. The 3-nearest neighbors are $(3, 5); (3, 8); (2, 11)$, after regression we have $y = 20 - 4.5x$, so y will be 5.6 when $x = 3.2$.

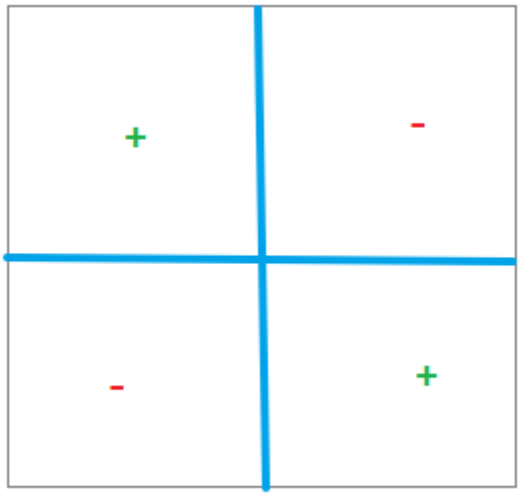
3. We have

$$(+1, +1) \to (+1, +1) : -1$$
$$(+1, -1) \to (+1, -1) : +1$$
$$(-1, +1) \to (-1, -1) : +1$$
$$(-1, -1) \to (-1, +1) : -1$$

That is



The separator is $x_1 x_2 = 0$, the margin is 1. And drawing it back to original space. It will be the two axises.



4.

$$\begin{aligned} \|\Phi(x_i) - \Phi(x_j)\| &= \sqrt{\Sigma_{k=1}^{D}(x_{i,k} - x_{j,k})^2} \\ &= \sqrt{\Sigma_{k=1}^{D}(x_{i,k} - x_{j,k}) \times (x_{i,k} - x_{j,k})} \\ &= \sqrt{K(x_i - x_j, x_i - x_j)} \end{aligned}$$

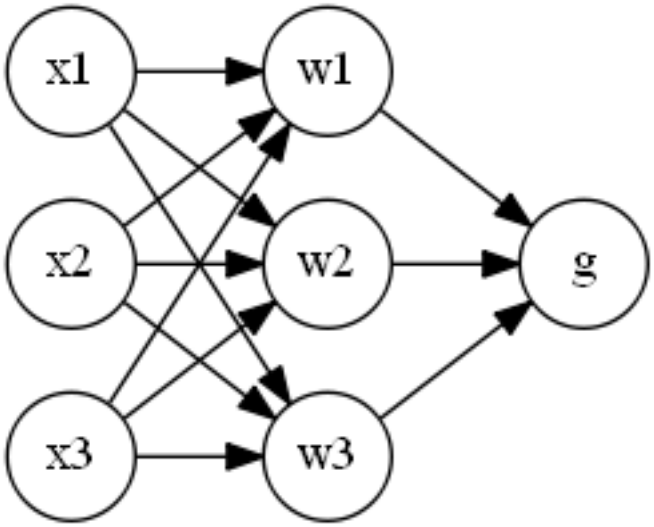
5. The true table will be

x_1	x_2	x_3	$XOR(AND(x_1, x_2), x_3)$
+	+	+	-
+	+	-	+
+	-	+	+
+	-	-	-
-	+	+	+
-	+	-	-
-	-	+	+
-	-	-	-

That is

$$g(\vec{x}) = XOR(AND(x_1, x_2), x_3) = ((\neg x_1 \wedge x_3) \vee (\neg x_2 \wedge x_3) \vee (x_1 \wedge x_2 \wedge \neg x_3))$$

So the neural network will be(1s are omitted)



Where

$$\begin{aligned} w_1 &= sign(-x_1 + x_3 - 1.5) \\ w_2 &= sign(-x_2 + x_3 - 1.5) \\ w_3 &= sign(x_1 + x_2 - x_3 - 2.5) \\ g &= sign(w_1 + w_2 + w_3 + 1.5) \end{aligned}$$