INDRAJEET
SARAVANAN
saravanan.8
@osu.edu

HOMEWORK - 3

CONCEPT CHECK

1. $P(B/+j,+m)$

By referring to slide 3 in Bayes' Net inference, we can rewrite the same as

$$= \sum_{e,a} P(B,e,a,+j,+m)$$

$$= \sum_{e,a} P(B)\cdot P(e)\cdot P(a/B,e)\, P(+j/a)\; P(+m/a)$$

$$= P(B)\cdot P(+e)\cdot P(+a/B,+e)\cdot P(+j/+a)\cdot P(+m/+a)$$
$$+ P(B)P(+e)P(-a/B,+e)\cdot P(+j/-a)\,P(+m/-a)$$
$$+ P(B)\cdot P(-e)\, P(+a/B,-e)\cdot P(+j/+a)\cdot P(+m/+a) +$$
$$P(B)\cdot P(-e)\cdot P(-a/B,-e)\cdot P(+j/-a)\cdot P(+m/-a)$$

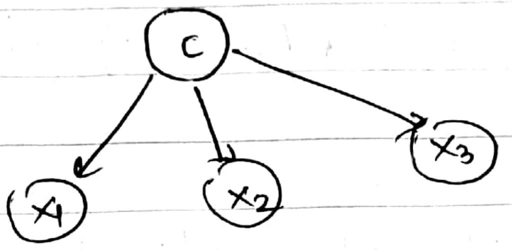
$$= (0.001)(0.002)(0.95)(0.9)(0.7) +$$
$$(0.001)(0.002)(0.05)(0.05)(0.01) +$$
$$(0.001)(0.998)(0.94)(0.9)(0.7) +$$
$$(0.001)(0.998)(0.06)(0.05)(0.01)$$

$$= 1.197\times10^{-6} + 5\times10^{-11} + 5.91\times10^{-4} +$$
$$2.994\times10^{-8}$$

$$\approx 5.92\times10^{-4}.$$

Q1.

2.



$P(C_1) = 0.3$

$P(C_2) = 0.5$

$P(C_3) = 0.2$

$P(x_1 = T / C_1) = 0.7$ $\quad$ $P(x_1 = F / C_1) = 0.3$

$P(x_1 = T / C_2) = 0.4$ $\quad$ $P(x_1 = F / C_2) = 0.6$

$P(x_1 = T / C_3) = \cancel{0.5}\ 0.2$ $\quad$ $P(x_1 = F / C_3) = 0.8$

$P(x_2 = T / C_1) = 0.9$ $\quad$ $P(x_2 = F / C_1) = 0.1$

$P(x_2 = T / C_2) = 0.5$ $\quad$ $P(x_2 = F / C_2) = 0.5$

$P(x_2 = T / C_3) = 0.7$ $\quad$ $P(x_2 = F / C_3) = 0.3$

$P(x_3 = T / C_1) = 0.6$ $\quad$ $P(x_3 = F / C_1) = 0.4$

$P(x_3 = T / C_2) = 0.4$ $\quad$ $P(x_3 = F / C_2) = 0.6$

$P(x_3 = T / C_3) = 0.2$ $\quad$ $P(x_3 = F / C_3) = 0.8$

$$P(C / x_1 = F, x_2 = T, x_3 = F) = \frac{P(x_1 = F, x_2 = T, x_3 = F / C) \cdot P(C)}{P(x_1 = F, x_2 = T, x_3 = F)}$$

$$\approx P(x_1 = F, x_2 = T, x_3 = F / C) \cdot P(C)$$

$$\approx P(x_1 = F/c) \cdot P(x_2 = T/c) \cdot P(x_3 = F/c) \cdot P(c)$$

when $c = c_1$;

$$P(x_1 = F/c_1)\, P(x_2 = T/\cancel{c_2})\, P(x_3 = F/c_1) \cdot P(c_1)$$

$$= 0.3 \times 0.9 \times 0.4 \times 0.3$$

$$S_1 = 0.0324.$$

when $c = c_2$

$$P(x_1 = F/c_2)\, P(x_2 = T/c_2)\, P(x_3 = F/c_2)\, P(c_2)$$

$$= 0.6 \times 0.5 \times 0.6 \times 0.5$$

$$S_2 = 0.09.$$

when $c = c_3$

$$P(x_1 = F/c_3) \cdot P(x_2 = T/c_3)\, P(x_3 = F/c_3) \cdot P(c_3)$$

$$= 0.8 \times 0.7 \times 0.8 \times 0.2$$

$$S_3 = 0.0896.$$

$$S_1 + S_2 + S_3 = 0.212.$$

After normalizing,

$$\frac{0.0324}{0.212}, \quad \frac{0.09}{0.212}, \quad \frac{0.0896}{0.212}.$$

The probabilty distribution is.

$\Rightarrow$ 0.1528, ~~0.41908~~ 0.42452, 0.4226.

4

3. sigmoid function is

$$s(z) = \frac{1}{1+e^{-z}}.$$

$$s'(z) = s(z)\left(1 - s(z)\right).$$

The saturation problem exists in sigmoid function where the $s'(z)$ tends to zero which in turn means slower training.
To solve this we can use a ReLU (Rectified linear unit) as activation function.
However, in ReLU we face the problem of dying ReLU. The neurons in the region of negative values won't make a difference.
Hence, leaky ReLU would be the ideal solution.

$$y = \begin{cases} ax & \text{for } x < 0 \\ x & \text{for } x \geq 0 \end{cases}.$$

a can be a decimal value.

FUN WITH PROOFS.

1. $O_5 = g\left(W_{3,5} \cdot g\left(W_{1,3} \cdot x_1 + W_{2,3} \cdot x_2 + B_3\right) + W_{4,5} \cdot g\left(W_{1,4} \cdot x_1 + W_{2,4} \cdot x_2 + B_4\right) + B_5\right).$

A multilayer perceptron can be represented as an affine function if activation function for each neuron is an affine function.

$$O_5 = \overline{W_1} \cdot x_1 + \overline{W_2} \cdot x_2 + \overline{W_3}$$

$$O_5 = W_{3,5} \left(W_{1,3} x_1 + W_{2,3} x_2 + B_3 \right) + W_{4,5} \left(W_{1,4} x_1 + W_{2,4} x_2 + B_4 \right) + B_5 .$$

$$O_5 = W_{3,5} W_{1,3} x_1 + W_{3,5} \cdot W_{2,3} \cdot x_2 + W_{3,5} B_3 + W_{4,5} \cdot W_{1,4} \cdot x_1 + W_{4,5} \cdot W_{2,4} \cdot x_2 + W_{4,5} \cdot B_4 + B_5 .$$

$$= \left(W_{3,5} \cdot W_{1,3} + W_{4,5} \cdot W_{1,4} \right) x_1 + \left(W_{3,5} \cdot W_{2,3} + W_{4,5} \cdot W_{2,4} \right) x_2 + \left(W_{3,5} + W_{4,5} + 1 \right) B .$$

$$O_5 = \overline{W_1} x_1 + \overline{W_2} \cdot x_2 + \overline{W_B}$$

Therefore, we can prove that it can be represented like the affine function.