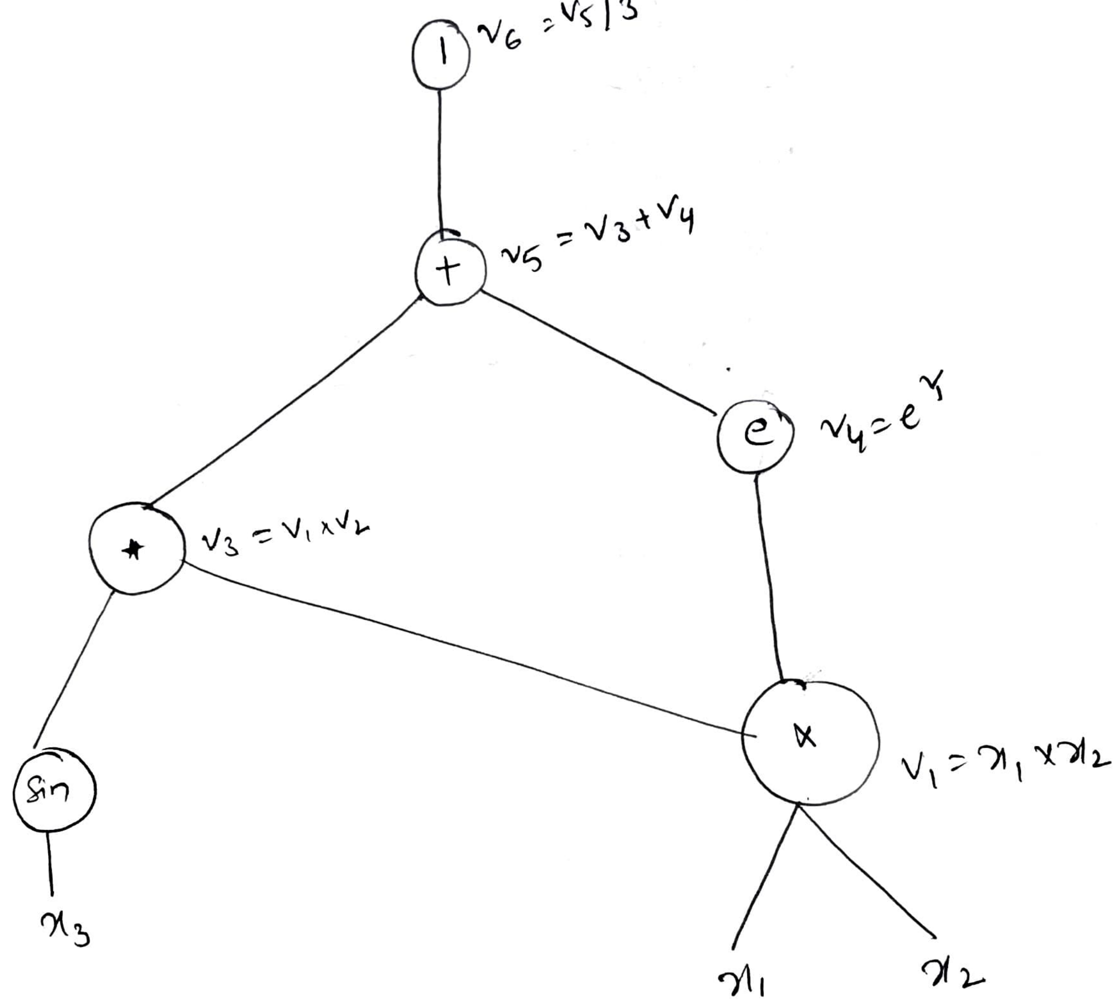
4)

$$f(x, y, z) = \frac{1}{3}\left(x_1 x_2 \sin(x_3) + \exp^{x_1 x_2}\right).$$

(a) Computation graph:-



(b) Computation ~~graph~~ trace :-

$v_1 = v_0 x_2$, $v_0 = x_1$

$v_2 = \sin(x_3)$

$v_3 = x_1 x_2 \sin(x_3) = \cancel{x} \, v_1 v_2$

$v_3 = v_1 v_2$

$v_4 = e^{x_1 x_2}$, $v_4 = e^{v_1}$

Since, $f(x, y, t) = \frac{v_3 + v_4}{3}$

$v_5 = v_3 + v_4$

$v_6 = v_5 / 3$

$f(x, y, t) = v_6$ //

(c) AD Trace for $x_1 = 1, x_2 = 0, x_3 = 0$

$(v_0, v_0') = v_0 = 0$, $v_0' = x_1'$

$(v_1, v_1') = (x_1 x_2 , x_1 x_2' + x_1' x_2)$

when $x_1 = 1, x_2 = 0, x_3 = 0$

$\rightarrow v_1 = 0$

$\rightarrow v_1' = x_2'$

$(v_2, v_2') = (\sin(x_3), \cos(x_3) \times x_3')$

$(\sin(x_3) , 1 \times x_3')$

$v_2' = x_3' \quad , v_2 = 0.$

$(v_3, v_3') = (v_1 v_2 \, , \, v_1 v_2' + v_1' v_2)$

$v_3 = 0$

$v_3' = 0$

$(v_4, v_4') = (e^{v_1}, e^{v_1} \times v_1')$

$v_4 = 1 \quad , \quad v_4' = x_2'$

$(v_5, v_5') = (v_3 + v_4 \, , \, (v_3 + v_4))$.

$v_5 = 1 \qquad v_5' = x_2'.$

$(v_6, v_6') = \left(\frac{v_5}{3} \, , \, \frac{1}{3} \times v_5'\right)$

$v_6 = \frac{1}{3} \quad , \quad v_6' = \frac{x_2'}{3}$

(d) Reverse AD Trace - $x = 1, y = 0, z = 0.$

$$v_6' = \frac{x_2'}{3}$$

$$\frac{x_2'}{3} = 1$$

$$x_2' = 3$$

$v_5' = x_2'$

$v_5' = 3$

$v_4' = x_2'$

$v_4 = 3$

$v_3' = 0$

$v_2' = 0$

$v_1' = 3$ //.

5)

(a)

ROC curve - works well for balanced datasets

Precision - recall curve - Imbalanced datasets

$$FPR = \frac{\text{False positive}}{\text{False positive} + \text{True negative}}$$

$$TPR(\text{Sensitivity}) = \frac{TP}{TP + FN}$$

TPR ↑ FP ↓

(a) We must prefer model A if the value of true positive in our classification problem is very important. For example, if we have a rain prediction model, we need to tell that when it will rain so that people are prepared. So, in our case even if we have a lower probability of raining we must classify it as rain will occur if not we cannot be prepared, in case if it did not rain it won't affect us much as people were just prepared for the rain.

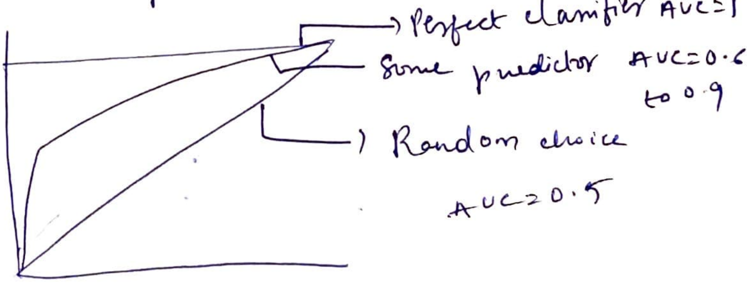
We need to prefer model B when we care about False positive rate. For example, if we have a model to predict. for cancer, suppose we classify someone without cancer as having cancer and we do chemotherapy to them based on our model predictions. It might become fatal to them in that case false positives are more important. So, it always depends on applications.

(b) In the low sensitivity range i.e, in the low false positive rate range Classifier A should be preferred. Like in the example above, we need to tell that when it will rain so that people are prepared even if the outcome turns to be false.

(c) We must prefer classifier B, in the high false positive rate range i.e or high sensitivity range. For example, in the cancer example, missing cases of actual cancer could lead to delays in treatment that negatively affect outcomes.

(d) Factors that determine AUC:-

↳ AUC depends on true positive and false positive rates.



So, if AUC < 0.5 → very bad ie, it is like a coin toss, prediction rate is 50%.

if 0.5 < AUC < 1 some predictive power.

AUC helps in choose which model is better than others.