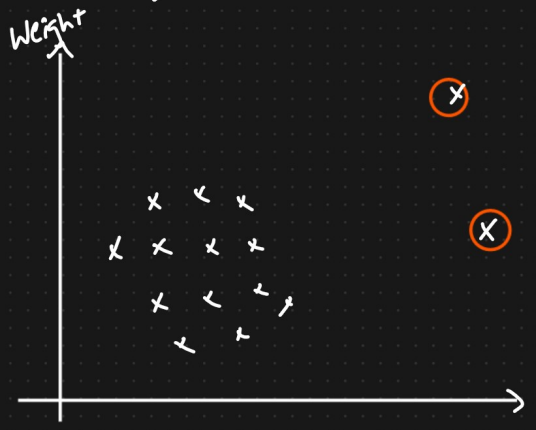
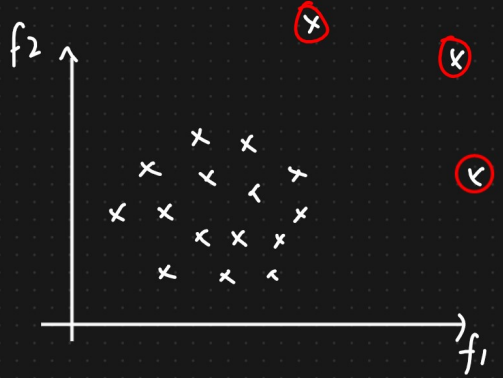
Anomaly Detection [To detect Outliers]

⇓ Important for a problem statement.

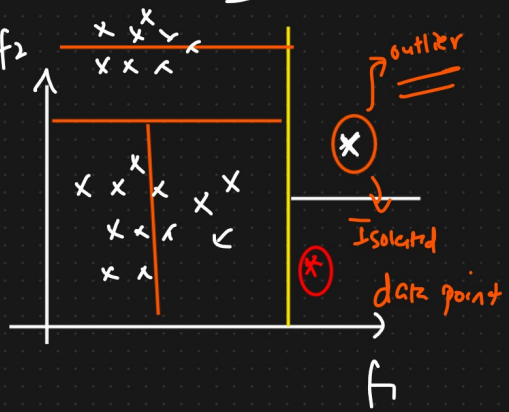


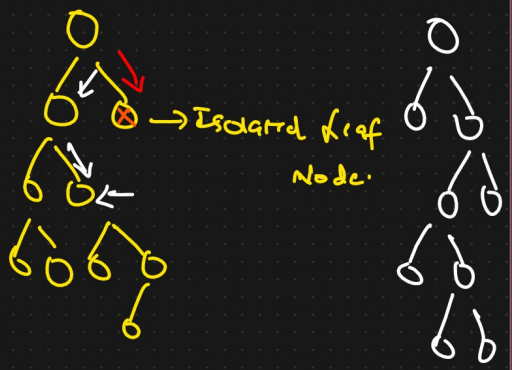


Anomaly Score | Isolation Trees

① Isolation Forest [Decision Trees]

f1	f2	f3	f4
-	-	-	-
-	-	-	-
-	-	-	-
-	-	-	-





Anomaly Score

Mathematical Formula : Compute anomaly Score for a new point.

$$S(x,m) = 2^{\frac{-E(h(x))}{c(m)}}$$

m = No. of Data points

x = Data point =>

E(h(x)) = Average search depth for x from the isolated tree.

c(m) = Average depth of all the data points

Threshold ≥ 0.5

$F(h(x)) << G(m) \Rightarrow S(x,m) \approx 1 \Rightarrow$ Anamoly score $\Rightarrow$ Outliers.

$F(h(x)) >> G(m) \Rightarrow S(x,m) \approx 0.5 \Rightarrow$ Normal data points