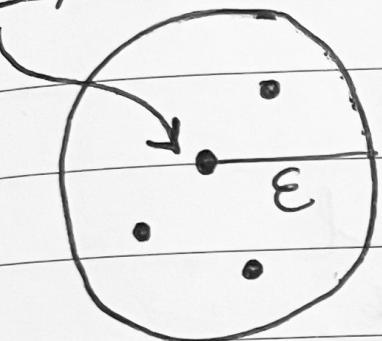


DBSCAN

ϵ & minpoint = 3

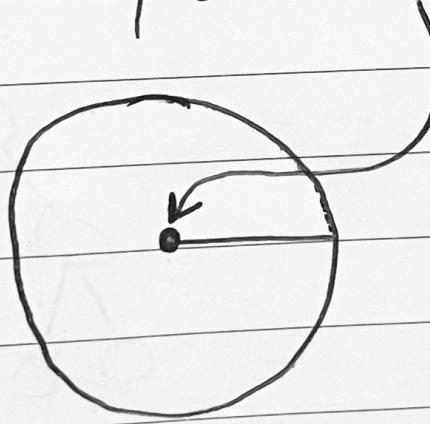
CP



BP



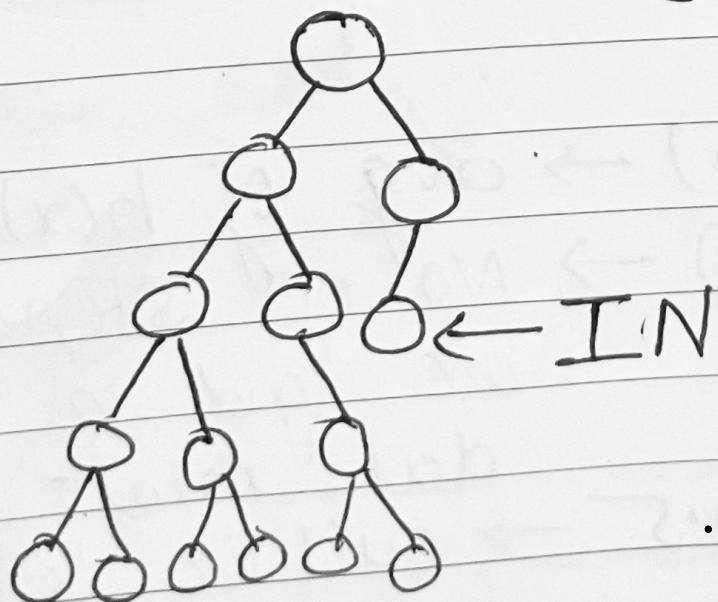
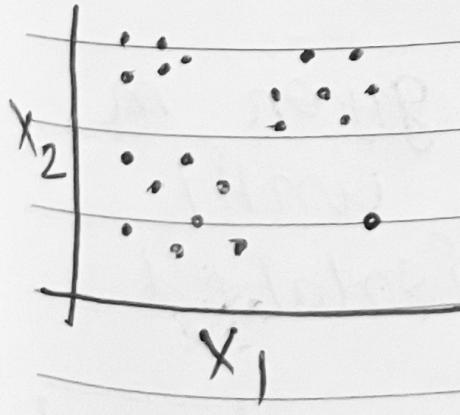
Noise/outlier



Isolation Forest

$$\{x_1 \quad x_2\}$$

make many ITs



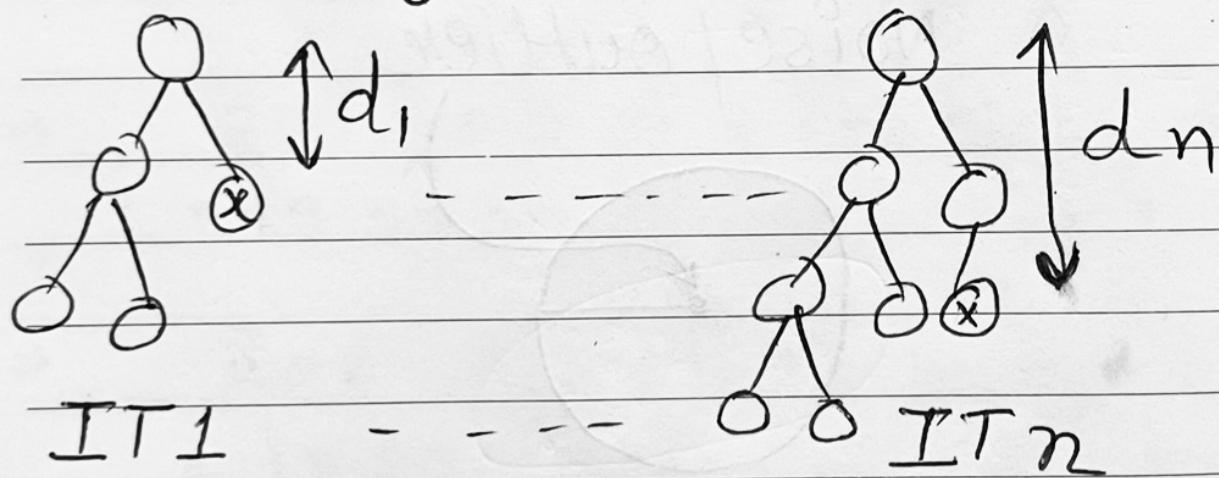
$$S(x, m) = 2 \frac{-E(h(x))}{c(m)}$$

$m \rightarrow$ size of data

$x \rightarrow$ single DP

↳ Normal
↳ Outlier

$E(h(x))$



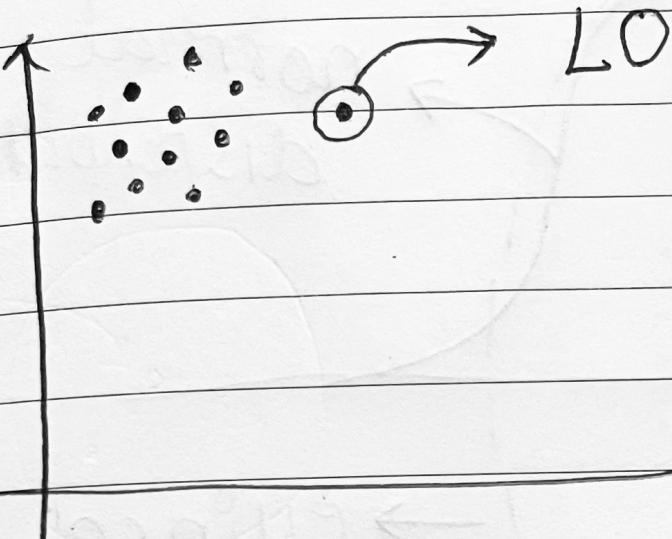
$$\underbrace{d_1 + d_2 + \dots + d_n}_n = E(h(x))$$

$c(m) \rightarrow$ avg of $h(x)$ given m

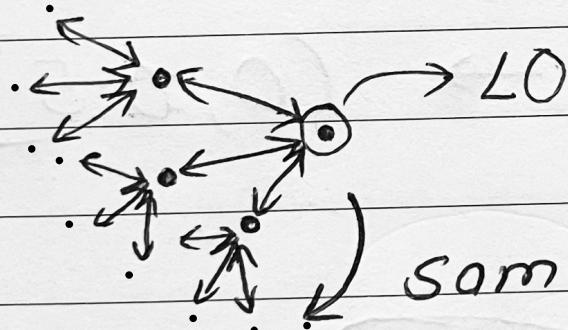
$h(x) \rightarrow$ no. of steps until we get a isolated data point

$S \geq 0.5 \rightarrow$ outlier else normal

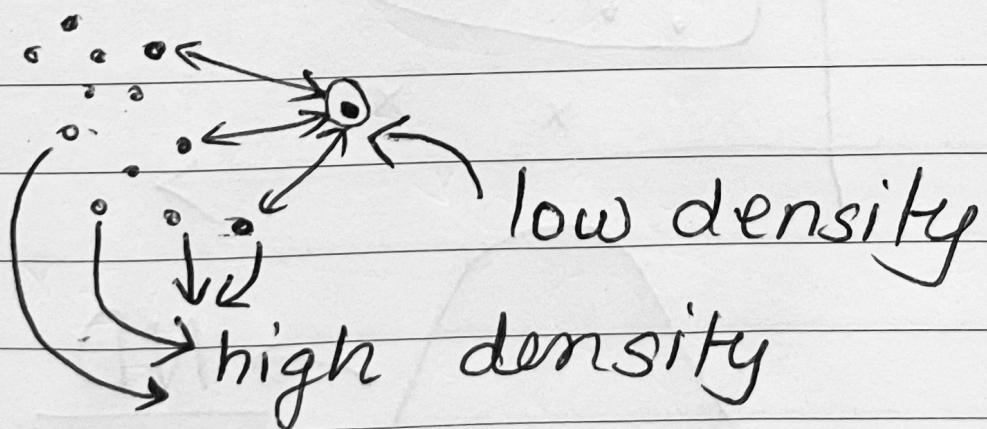
LOF



for $k = 3$



some for its
neighbors



Contamination $\rightarrow [0, 0.5]$

