05 June 2025 21:29

Day-2

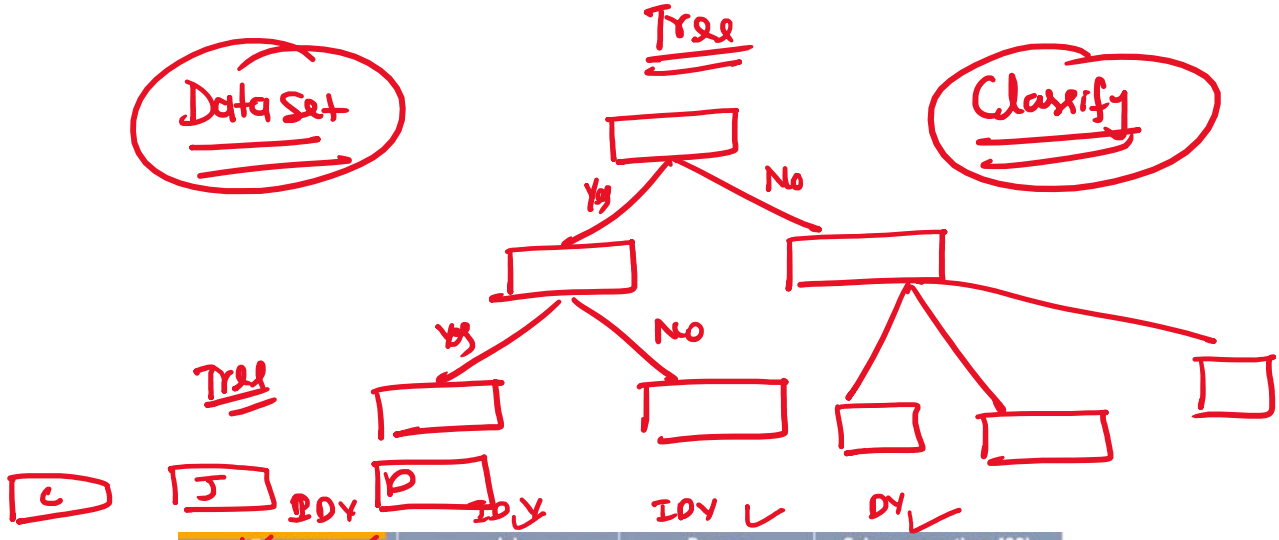
Recap Day-1

- ML
- AI, ML, DL, Non-AI
- Supervised Learning
- Un-supervised Learning
- Data Splitting ← Train, Test, validate
- ML life cycle

Animated Quiz

Continuous
- Linear Regression ← Bi-variant (x, y) ; y = mx+c
  - Multi-variant (x_1, x_2, y)

Categorical
- Logistic Regression → sigmoid



C J ID

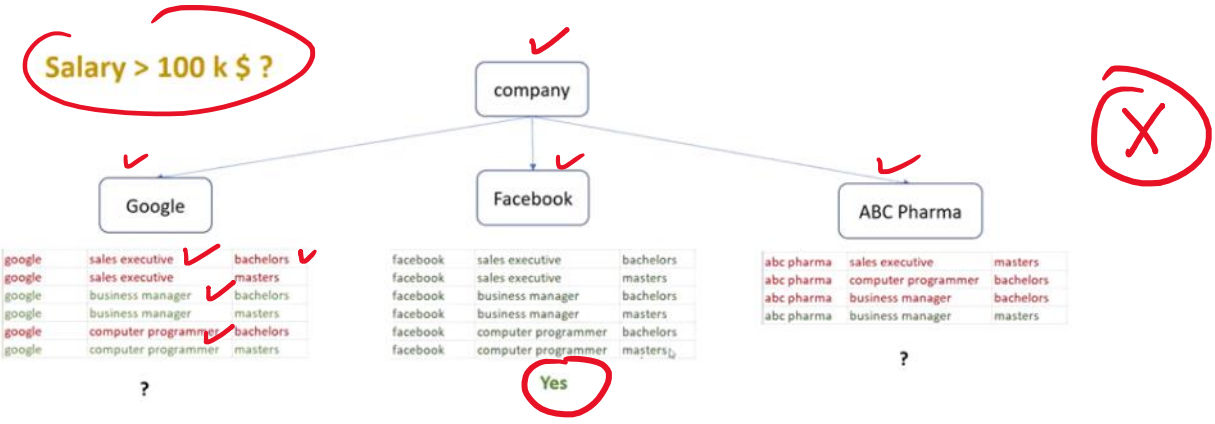
IDV IDV IDV DV

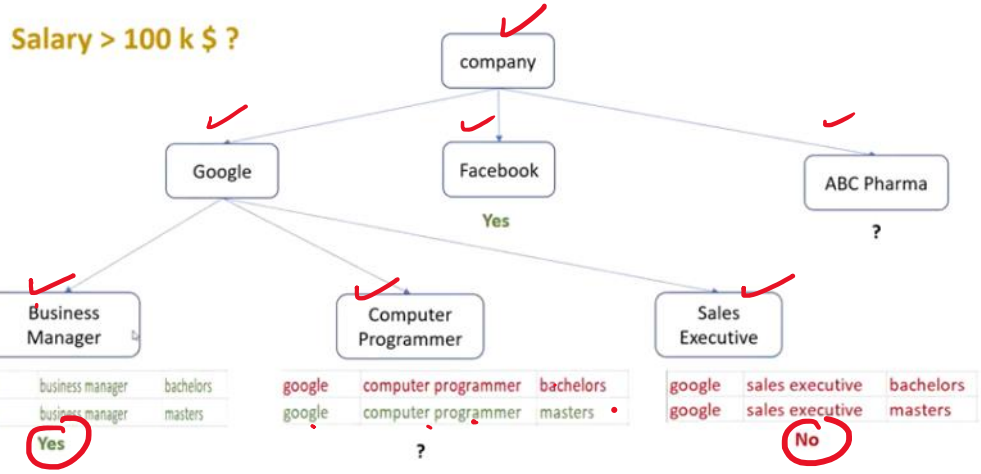
company	Job	Degree	Salary_more_then_100k
google	sales executive	bachelors	0
google	sales executive	masters	0
google	business manager	bachelors	1
google	business manager	masters	1
google	computer programmer	bachelors	0
google	computer programmer	masters	1
abc pharma	sales executive	masters	0
abc pharma	computer programmer	bachelors	0
abc pharma	business manager	bachelors	0
abc pharma	business manager	masters	1
facebook	sales executive	bachelors	1
facebook	sales executive	masters	1
facebook	business manager	bachelors	1
facebook	business manager	masters	1
facebook	computer programmer	bachelors	1
facebook	computer programmer	masters	1

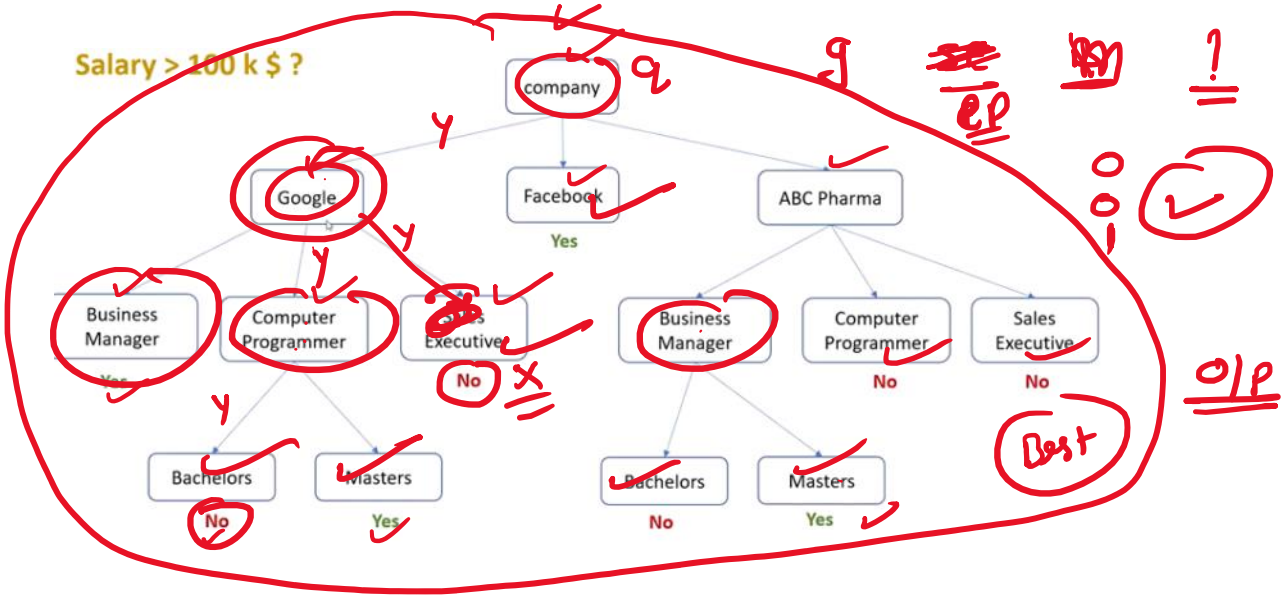
100k $/year

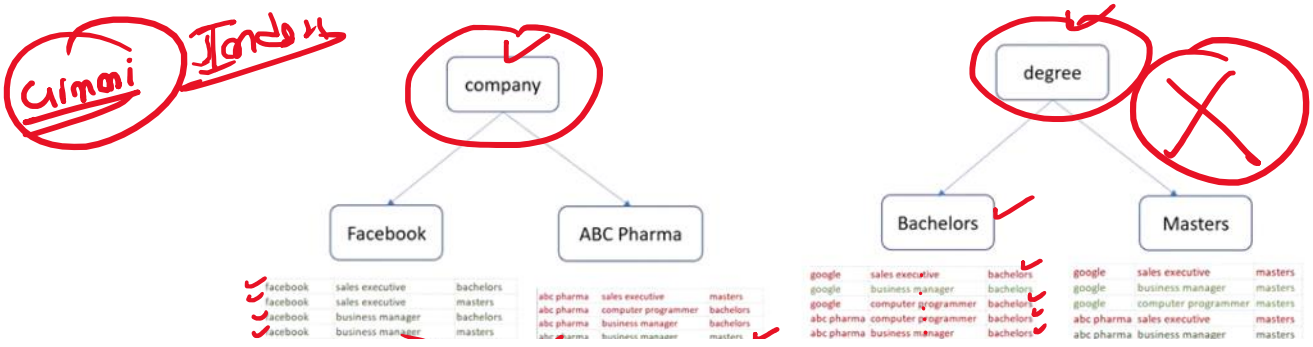
Prob

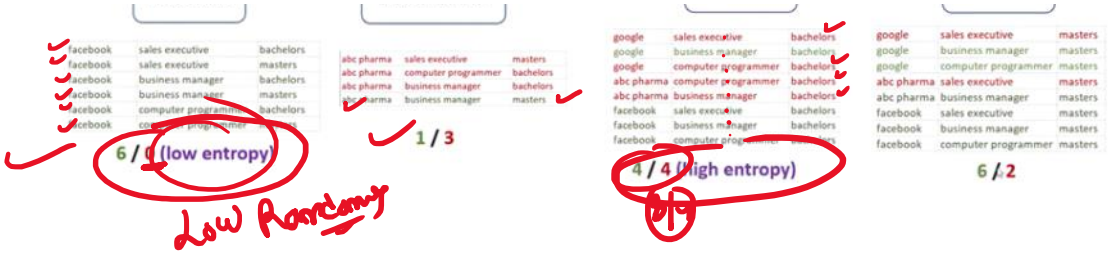
g SE b ;

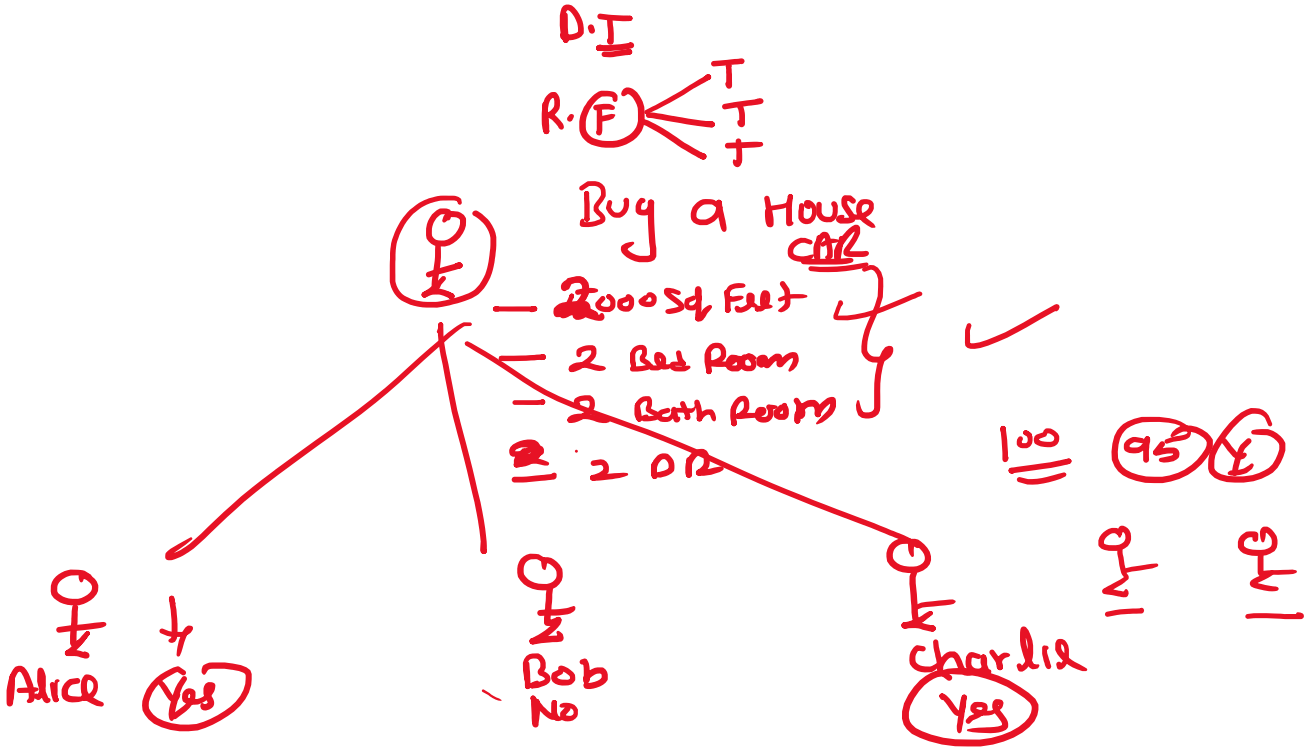

Salary > 100 k $?
company
Google
Facebook
ABC Pharma
google sales executive bachelors
google sales executive masters
google business manager bachelors
google business manager masters
google computer programmer bachelors
google computer programmer masters
facebook sales executive bachelors
facebook sales executive masters
facebook business manager bachelors
facebook business manager masters
facebook computer programmer bachelors
facebook computer programmer masters
abc pharma sales executive masters
abc pharma computer programmer bachelors
abc pharma business manager bachelors
abc pharma business manager masters
?
Yes
?

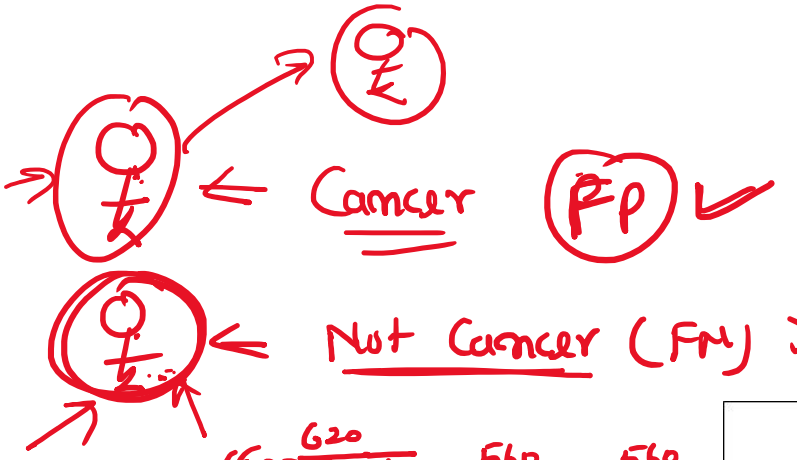

Salary > 100 k $?
company
Google
Facebook
ABC Pharma
Yes
?
Business Manager
Computer Programmer
Sales Executive
google business manager bachelors
google business manager masters
google computer programmer bachelors
google computer programmer masters
google sales executive bachelors
google sales executive masters
Yes
?
No


Salary > 100 k $?
company
Google
Facebook
ABC Pharma
Yes
Business Manager
Computer Programmer
Executive
No
Business Manager
Computer Programmer
Sales Executive
No
No
Bachelors
Masters
No
Yes
Bachelors
Masters
No
Yes


company
Facebook
ABC Pharma
degree
Bachelors
Masters

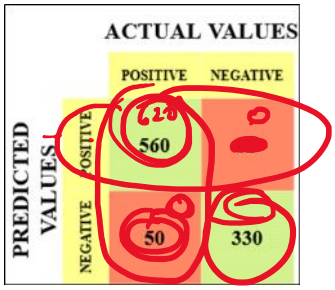






Precision = TP / (TP+FP)

Recall= TP / (TP+FN)



Recall: Recall provides us the information about the proportion of actual positives that was identified correctly.

Recall= TP / (TP+FN)

F1- Score: The F1 score is interpreted as a weighted average of the precision and recall, where F1 score of value 1 is best and score of value 0 is worst.

F1= 2 * [(Recall * Precision) / (Recall + Precision)]

Accuracy: Accuracy is a metric for evaluating classification models. Accuracy gives the information about the fraction of predictions that our model got right. In percentage;

TP
TN
FP
FN

TP = 560
FP = 60
FN = 50

F1= 2 * [(Recall * Precision) / (Recall + Precision)]

Accuracy: Accuracy is a metric for evaluating classification models. Accuracy gives the information about the fraction of predictions that our model got right. In percentage;

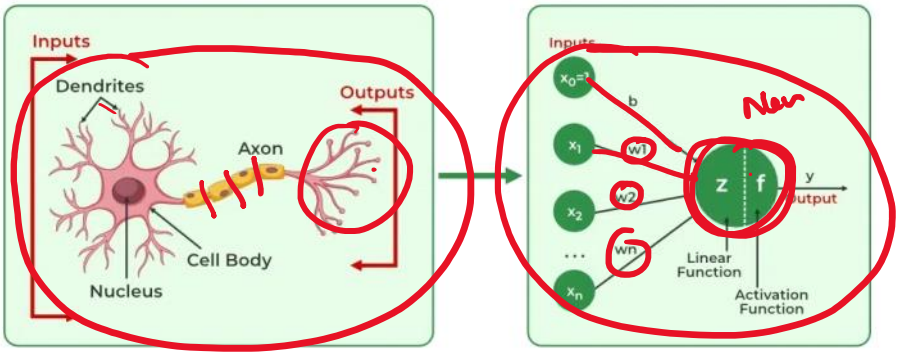
Accuracy = 100 * [(TP+TN) / (TP+TN+FP+FN)]

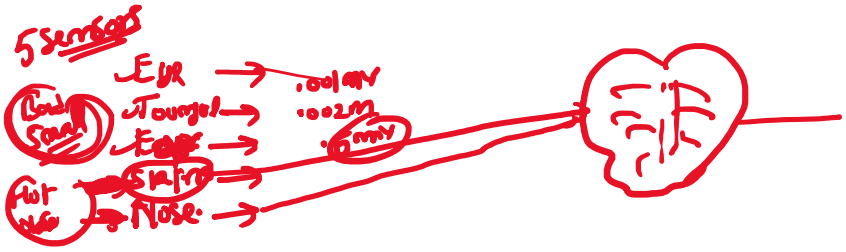
Here, TP= True Positive, TN= True Negative, FP = False Positive and FN = False Negative.

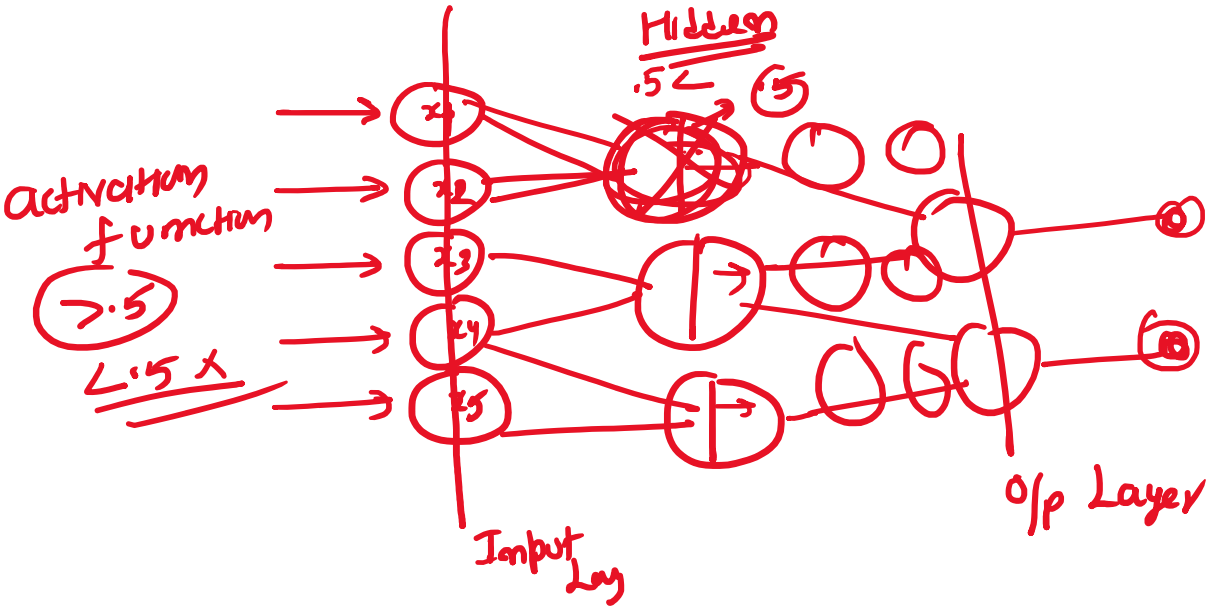
FP = 60
FN = 50
TN = 330

$F1 = 2 \times (1 \times 1)/(1+1) = \frac{2 \times 1}{2}$ -(1)

Kaggle







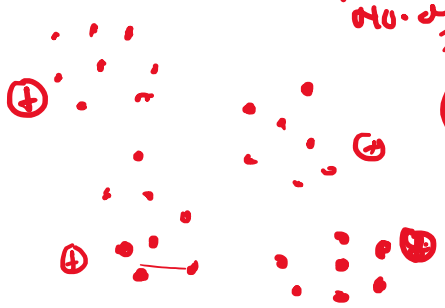
Unsupervised Learning

Clustering Algorithm

Elbow Method

- K means Algorithm

↑ No. of Clusters



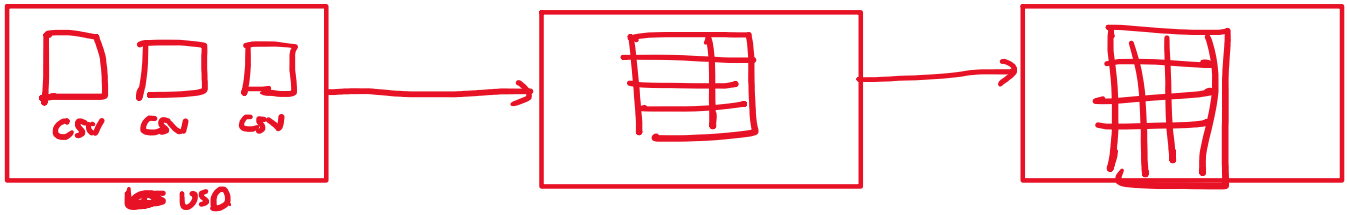
4

Centroid

E
Extraction

T
Transformation

L
Loading



Log

390x70