DOMS | Page No.
Date / /

Test AI & ML

1. Small K with Noisy Data
2. Small changes in Data lead to Different trees.
3. By Reducing Variance
4. All features are considered at each split
5. Target variable is categorical
6. Sigmoid
7. Accuracy
8. ~~Overfitting~~ Training time
9. Because distance calculation depend on scale
10. Logistic Regression

Ans) When tree needs to grow deep. then the Depth As Parameter is used in Overfitting in Decision tree. A tree learn all the training data and memorize it. They can learn noisy data as well. They have many nodes and branches because tree grow deep.

They can't memorize new data because they already read and memorize training data.

Ans 2 Bagging Address this problem by creating multiple trees on some training data

Random forest Address this problem by Adding layer on Bagging Randomness.

A3 Random forest working in detail:-

- Bootstrap sampling:- Boot strap sampling is also known as 'Bagging'. They create multiple trees of training data to reduce overfitting on in decision tree.

- Random feature selection:- for each tree which is created by bootstrap sampling: a random subset of feature is select.

- Majority voting:- In majority voting this resume find result or Prediction of Data.

DOMS | Page No.
Date / /

Fraud Detection Model

	Predicted Fraud	Predicted Not Fraud
Actual fraud	120	30
Actual Not fraud	50	800

- Calculate Accuracy
- " " Precision
- " " Recall
- " " F1 Score
- Is this model Acceptable for fraud Detection?

TP = 120
FN = 30
FP = 50
TN = 800
1000

① Accuracy :- $A = \frac{TP + TN}{Total} = \frac{120 + 800}{1000} = \frac{920}{1000}$

0.92 %.

② Precision :- $P = \frac{TP}{TP + FP} = \frac{120}{120 + 50} = \frac{120}{170} = 0.70\%$

③ Recall $= \frac{TP}{TP + FN} = \frac{120}{120 + 30} = \frac{120}{150} = 0.8\ \%$

④ F1 Score $= \frac{2 * Precision * Recall}{Precision + Recall}$

$F1 = \frac{2 \times 0.70 \times 0.8}{0.70 + 0.8} = \frac{1.12}{1.5} = 0.74$

DOMS | Page No.

Date / /

- Yes model is acceptable for fraud detection.