Interview questions.

① What is the difference between precision and recall?

(*) Precision :-

* Precision measures how many of the predicted positive cases are actually positive. It reflects the accuracy of positive Prediction.

Formula :-

$$\text{Precision} = \frac{TP}{TP + FP}$$

TP : True positive

FP : False positive

For ex :- In medical diagnostics for a serious disease, a false positive could lead to unnecessary anxiety or invasive tests.

⊛ Recall

* Recall measures how many of the actual positive cases are correctly identified, It reflects the model's ability to detect positive cases.

Formula

$$\text{Recall} = \frac{TP}{TP + FN}$$

FN: False Negative.

For ex:-

In detecting cancer a False Negative is far worse than False Positive.

⊛ Precision vs Recall Trade-off

* Increase precision often decreases recall.
* F1-Score is metric that balances Both

Formula.

$$F_1\text{-Score} = \frac{\text{Precion} \times \text{Recall}}{\text{Precision} + \text{Recall}}.$$

(7) What is cross-validation, and why is it important in binary classification?

Cross-validation is a statistical technique used to evaluate the performance of ML model.
A split the multiple subsets.

How does cross-validation work:-

1) Spliting dataset
2) Testing validation
3) Repeat validation testing
4) Average the result.

Types of cross validation:-

1) k-fold cross-validation
2) Statisfied k-fold
3) Leave one out (LOO) one of testing.

* why is cross-validation important in binary classification.

1) Prevents overfitting:-

Ensure the model is not too

specialized to the training data and perform well on new data.

2) Improve model Reliability :-

Evaluate performance across multiple subsets.

3) Handle Data imbalance :-

Stratified cross validation help's maintain the class balance

4) Maximize Data usage.

All data points are used for both training and validation making most of small datasets.

5) Reduces Bias & variance :-

Average performance across folds to provide a more balanced evaluation.