# TECHNO MAIN SALT LAKE

( FORMERLY TECHNO INDIA, SALT LAKE )

Name... Shohan Nayak.....

Roll No. 13030822138 ..... Stream..... CSE (AIML) .....

Subject Application of Machine Learning in industries (PEC-AIML601) Semester..... 6th .....

Invigilator's Signature ..... Date..... 20/3/2025 .....

(Part A)

1) Classification of Regression are the two most common supervised tasks.

2) To check the accuracy of the prediction of a validation set.

3) There are two model parameters are there in a linear regression problem with a single feature variable.

4) AUC value of a perfect classifier is 90-100%.

5) Precision is more important for a spam email detection system.

[part-B]

6) train-test-split is a method available in sklearn that we used to split the data into training set and testing set. we use 30% of the total data.

underfitting and overfitting :-

overfitting is a phenomenon where the machine learning model predicts too well on the training data that is seen data. that does not work well on unseen data that is testing data.

where underfitting happens when the model performs badly on the training data that is the seen data that predicts badly too well on the unseen data that is testing dataset.

prevention techniques :-

1) use of suitable model
2) scale the data.

(3) reduce the noice from the data

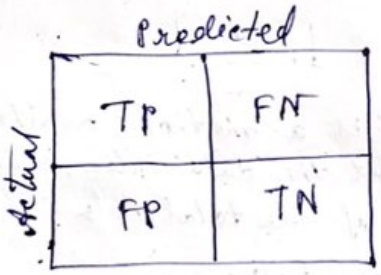
(4) Cross - validition.

7) bias → Bias is a terminology in machine learning. it signifies how a machine learning model predicting poorly in the training phase.
Basically the error in the training phase called as bias.

variance → varience is where the model gives error in the testing phase.

- To reduce it we take data that is properly encoded, scalled and the size of the data is moderate.
- bias-varriance tradeoff is a senario where if the model perform poorly on the training phase.

9) confusion matrix is the insight of model performance. It include correctly predicted and incorrectly predicted data value.




- TN = 82, FP = 3, FN = 5, TP = 10

Precision $= \frac{10}{10+3} = \frac{10}{13}$

recall $= \frac{10}{10+5} = \frac{10}{15}$

false negative rate $= \frac{5}{15} = \frac{1}{3}$

false positive rate $= \frac{3}{82+3} = \frac{3}{85}$.