Deep Salunkhe.
21102A0012.
BE CMPN A

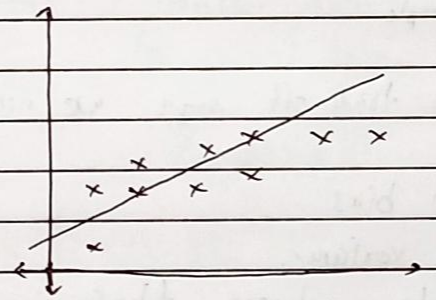
D D M M Y Y Y Y

Assignment 2

Q1] Explain the bias-variance trade-off in the context of model complexity. How does this trade-off influence the choice of model in real-world ML application?

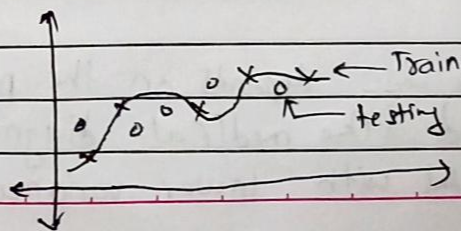
⇒ **Bias :**

→ When the model is simple, the modle might not fit the train set data point.
→ Here the difference between Actual and predicted value for train set is high this is bias.

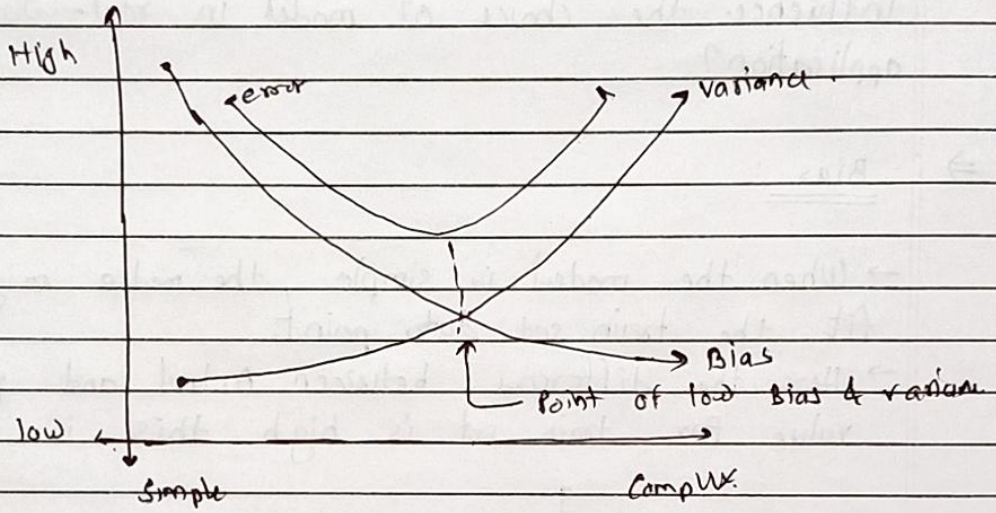


Variance :

→ When the model is too complex, the mode over fits the training set.
→ As a result the diff between the actual and predicted is low for the train set.
→ But the diff is high for other data set. this is variance



Bias-Variance Trade off.



Bias variance trade off says we need model that give

1. low bias
2. low variance.

The method to achieve this

* Regularization
* Boosting.
* Baggy

→ In real life if we have limited data we might choose a simpler model to avoid overfitting.

→ With large dataset, we can afford more complex model that can capture intricate patterns

→ The choice also depends on the problem domain. In some field like medical diagnosis we might prefer model with lower variance to ensure consistency

Q2) Imagine you are tasked with building an ML model to predict customer churn for a subscription-based service. Describe how you would handle the following aspects:

⇒

Choosing the right type of ML algo:

We can consider the following

- logistic regression: simple, interpretable, good for linearly seperated data.
- Random forest: Handles non-linear relationship, good for mixed data.
- Neural network: If we have large dataset with complex pattern

Choice depends on factors like data size, feature type, interpretability requirement.

Addressing potential ethical issues

- Ensure data privacy and compilance
- check for bias in the training data that could lead to unfair prediction
- Ensure transperancy.

Managing overfitting and underfitting

- Use cross validation on unseen data.
- Apply regularization technique to avoid overfitting

- Use ensemble methods to reduce both variance & bias
- monitor training and validation errors to detect overfitting early

Balancing the bia-variance trade off

- start with simple model and gradually increase complexity while monitoring performance.
- Use techniques like. pruning for decision trees to reduce variance
- Use techniques like bagging and boosting

Ensure the model. generalizes well to unseen data

- Use seperate test set that's not used. during model development.
- perform k-fold-cross validation to get a. robust estimation of model performance.
- Regularly retrain and validate the model on new data to account for concept drift.