Aayush Srivastava CM PN-A
21102A0008

ML Assignment-2

Q1

The bias variance tradeoff balances two types of prediction error

bias: Error from simplistic models missing data patterns, causing underfitting

Variance: Error from complex models capturing noise, causing overfitting.

Influence on Model choice:

Simple models ⇒ High bias, low variance
eg:- linear regression may miss patterns

Complex models ⇒ low bias, High variance
eg: DNN may overfit with small data

Real world ex,
For predicting customer churn

- logistic regression:- simple, interpretable,
- Random forest:- captures complex patterns, reduces bias
- Balanced approach:- start simple, increase complexity if validation data supports it.

Q2

Predicting customer churn:

1 choosing the Algo

- start with regression for interpretability
- Consider random forest or gradient boosting for performance

2 Addressing Ethical issues.

- Ensure Fairness and transparency
- Protect data and comply with regulations

3 Managing overfitting and underfitting

- Use cross validation
- Adjust model complexity based on performance

4 Balancing tradeoff

- Optimize using Validation data.
- start simple, increase as needed

5 Ensuring Generalization

- Use K fold cross validation
- Maintain a separate test set
- Monitor and retrain the model regularly.