STAGE4_TEAM_REPORT

1) Discuss bias versus variance tradeoff.

- Generally, this Bias vs variance tradeoff will give us information about prediction errors and helps to find out model accuracy.
- We can have better understanding regarding the underfitting and overfitting problems.
- By default, the linear regression model is highly biased and has low variance. With increasing the degree of polynomial regression, the bias will be reduced, and the variance will start increasing. For higher degrees we can observe the overfitting problem.
- So, if we need to select a model which is best then we need to select the model which has moderate bias and moderate variance.
- If we observe the linear regression model predicted trend line the root mean squared error is high. As the degree of polynomial regression model is increased the root mean squared error will keep on reducing.

2) Describe the trends as compared to other countries.

- The new cases and new deaths across USA are more compared to the other selected countries Indonesia, Nigeria, and Pakistan. But the peaks months of new cases and new deaths across the USA matches with the three selected countries. There might be huge difference in number of cases and deaths registered but the months are almost same.
- For new cases across USA the best model predicted trend line is non-linear regression model with degree 3 and if we observe for higher degrees 4 and 5 the trend line seems to be overfitted. For new deaths across USA the best model predicted trend line is also non-linear regression model with degree 3.
- For new cases across Indonesia the best model predicted trend line is non-linear regression model with degree 3 and if we observe for higher degrees 4 and 5 the trend line seems to be overfitted in this case too. For new deaths across Indonesia the best model predicted trend line is also non-linear regression model with degree 3.
- For new cases across Nigeria the best model predicted trend line is non-linear regression model with degree 4 and if we observe for higher degree 5 the trend line seems to be overfitted in this case. For new deaths across Nigeria the best model predicted trend line is also non-linear regression model with degree 3.
- For new cases across Pakistan the best model predicted trend line is non-linear regression model with degree 2 and if we observe for higher degrees 3,4 and 5 the trend line seems to be overfitted in this case. For new deaths across Pakistan the best model predicted trend line is also non-linear regression model with degree 3.