1. Model building steps
2. Outlier – notebook ( [link](https://towardsdatascience.com/ways-to-detect-and-remove-the-outliers-404d16608dba)) - Done
3. Missing value – Done
4. Training testing event rate
5. Linear regression
6. Assumptions
7. R square formula and adjusted r squared
8. Classification
9. Logistic regression explained
10. Roc AUC threshold decide
11. Confusion matrix
12. Multicollinearity
13. Bias variance
14. Overfitting underfitting
15. Sampling
16. Random forest
17. Bagging and boosting basic
18. SVM basic and kernel bsed method(Polynomial RBF(
19. K means and elbow method
20. Distance measures
21. All challenges
22. KNN basic in challenge document
23. Topic modelling steps and codes
24. LDA and beta distribution
25. P value, mean median sd variance hypothesis testing all tests (T z chi square)
26. Variable importance
27. Decision tree gini etc
28. Sentiment analysis
29. PCA and it works using eigen values etc

**Coding**

**Classification python code we have developed**

**Concat merger lambda function apply python code**

**In R all apply and joins**

**Text mining code in R**

**Python data types(list tuple etc)**