**USING DENTAL METRICS TO PREDICT GENDER**

1. Import the necessary libraries such as pandas, NumPy, matplotlib, seaborn, sklearn.
2. Using minmax\_scale from sklearn.preprocessing to standardize the features in the data set.
3. Using confusion matrix imported from sklearn.metrics for evaluation.
4. Importing GridSearchCV to determine optimal parameters. Train\_test\_split for training and testing the data against the model used. Both imported from sklearn.model\_selection.
5. Read the Dentistry Dataset.csv file to load the data and store it in a variable named dataset.
6. Dataset.info() helps to understand the various features, the data type and the non-null count, range index and the memory usage.
7. Dataset.head() command gives the number of entries in the dataset. Default is 5.
8. Data pre-processing is carried out by counting the total number of null entries in the data set for every column.
9. Statistical analysis to understand mean, min, max, standard deviation and other parameters using the dataset.describe().
10. Visualizing different canine distances across values by plotting line chart and palette colors.
11. Encoding the gender column from categorical to numerical by allotting female as and male as 1 and mapping them to the Gender column.
12. Splitting the data between independent and dependent variables. The independent variables are stored in X while the dependent ones are stored in y.
13. Split the data into test and train using train\_test\_split module.
14. Loding different modules such as Logistic Regression, Decision Tree classifier, Random Forest Classifier, and XGBoost Classifier.
15. Using Logistic Regression to calculate the accuracy along with confusion matrix, roc curve, auc curve.
16. Using # Decision Tree Classifier with confusion matrix, roc curve, auc curve.
17. Using Random Forest Classifier with confusion matrix, roc curve, auc curve
18. Using Gradient Boosting Classifier with confusion matrix, roc curve, auc curve
19. **Accuracy of Logistic Regression is 92.27 %**
20. **Accuracy of Decision Tree Classifier is 100 %**
21. **Accuracy of Random Forest Classifier is 99.09 %**
22. **Accuracy of Gradient Boosting Classifier is 100 %**