# Tooth Measurement Dataset

## Model Building:

Let us build model using Logistic Regression, K-Nearest Neighbours and Stochastic Gradient Descent Machine Learning Algorithms.

## Feature Selection:

For good predictions of the outcome, it is essential to include the good independent variables (features) for fitting the model (e.g., variables that are not highly correlated). If we include all features, there are chances that we may not get all significant predictors in the model.

we visualize how logistic regression model uses the different features and which features have greater effect.

Left canine width intraoral, left canine width casts, Left Canine index intraoral, left canine index Casts and Inter canine Distance Intraoral have significant influence on the model.

Inter canine distance casts, right canine index casts, Right canine index intraoral have negative influence on the prediction.

Chart, bar chart

Description automatically generated

The idea of repeatedly constructing a model and choosing either the best or worst performing feature, setting the feature aside and then repeating the process with the rest of the features. This process is applied until all features in the dataset are exhausted and check P values of the independent features and we filter down on few of the features.

We analyse various performance measures like accuracies, confusion matrix, sensitivity, precision etc on all the below models with different features as input.