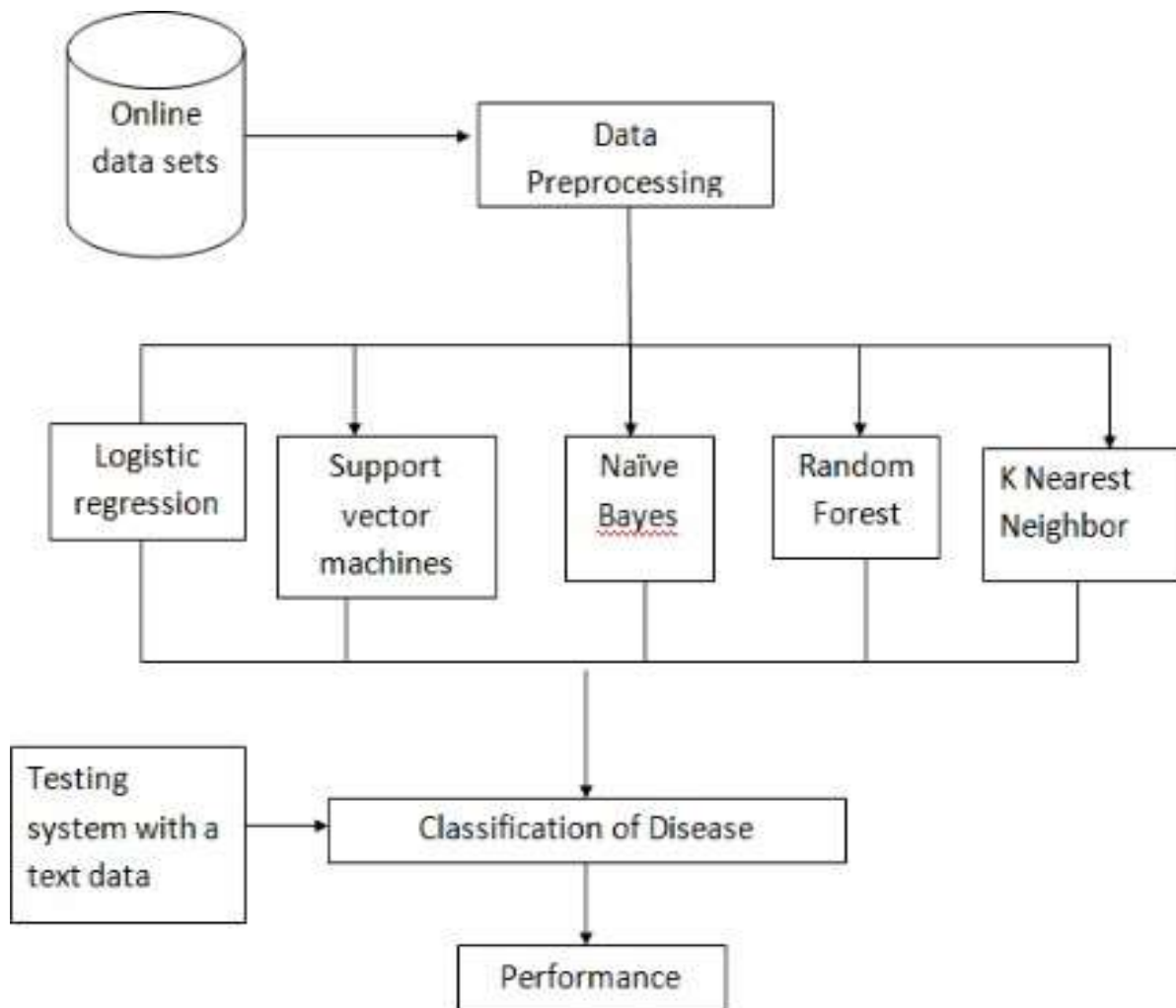# Statistical Machine Learning Approaches to Liver Disease Prediction

TEAM ID: PNT2022TMID42133

Team Leader : T.Ramesh
Team Member: G.Karunakaran
Team Member: A.Deepak
Team Member: P.Hariprasath

Project Workflow

```
┌─────────────┐
│   Online    │────────▶  ┌──────────────────┐
│  data sets  │           │      Data        │
└─────────────┘           │  Preprocessing   │
                          └──────────────────┘
```

Logistic regression	Support vector machines	Naïve Bayes	Random Forest	K Nearest Neighbor

```
┌──────────────────┐           ┌──────────────────────────┐
│     Testing      │────────▶  │  Classification of Disease │
│  system with a   │           └──────────────────────────┘
│    text data     │
└──────────────────┘                        │
                                             ▼
                                   ┌──────────────────┐
                                   │   Performance    │
                                   └──────────────────┘
```

Serial Number	Attribute
1	Age
2	Gender
3	Total Billirubin
4	Direct Bilirubin
5	Alkaline Phosphotase
6	Alamine Aminotransferase
7	Total Protiens
8	Albumin
9	Albumin and Globulin Ratio
10	Data Set Class

The performance evaluation of various machine learning algorithms like correctly classified instances, incorrectly classified instances, kappa statistic, Mean absolute error (MAE), Root Mean square error (RMSE)are to be

discussed. We are about to do calculation of True positive rate, False positive rate Precision, Recall, F-Measure, Roc values and confusion matrix of various considered machine learning algorithms.