Statistical Machine Learning Approaches to Liver Disease Prediction

TEAM ID: PNT2022TMID51504
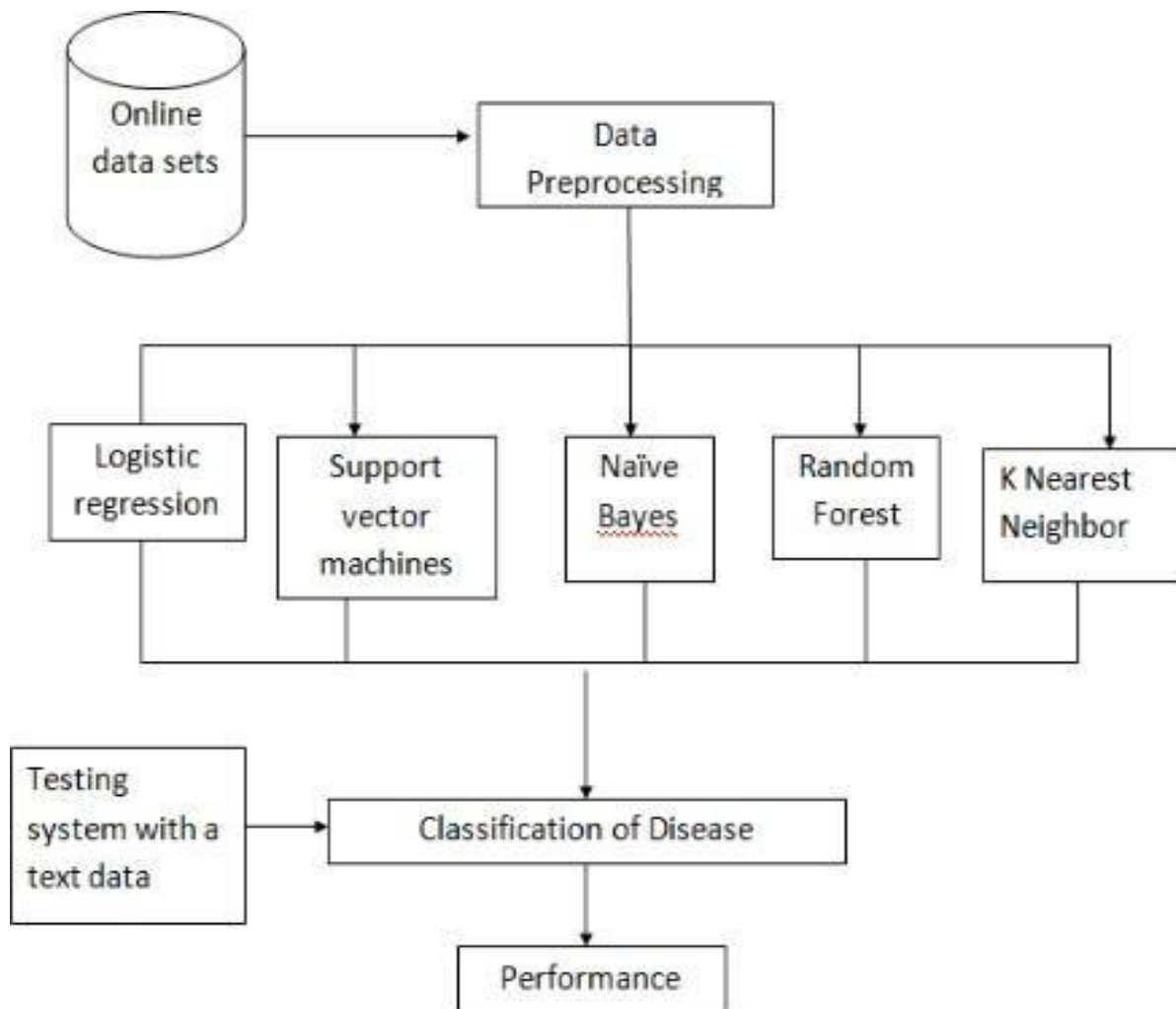
Team leader: L.G.Abitha

Team Membe: S.sahaya Stephitha

Team Member : Sajitha.T

Team Member : S.Sowmiya

Project Workflow


```
                    ┌─────────────────┐
   ╔═══════════╗    │      Data       │
   ║  Online   ║───►│ Preprocessing   │
   ║ data sets ║    └─────────────────┘
   ╚═══════════╝             │
                             ▼
 ┌──────────┬──────────┬──────────┬──────────┬──────────┐
 │ Logistic │ Support  │  Naïve   │ Random   │K Nearest │
 │regression│ vector   │  Bayes   │ Forest   │Neighbor  │
 │          │ machines │          │          │          │
 └──────────┴──────────┴──────────┴──────────┴──────────┘
                             │
 ┌──────────┐                ▼
 │ Testing  │    ┌───────────────────────────┐
 │system with a──►│ Classification of Disease │
 │ 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.