CLASSIFICATION ASSIGNMENT

- We have to predict the Chronic Kidney Disease based on several parameters
- In this data, the prediction is clear so, it should be considered as a Supervised Learning. The domain is Machine Learning
- Total number of Rows: 400 and Total number of Columns:25
- Here, we are using **One Hot encoding** to convert the String into Nominal data "dataset=pd.get_dummies(dataset,dtype=int,drop_first=True)
  dataset"

CLASSIFICATION REPORT:-

- **RANDOM FOREST**

The classification report:

	precision	recall	f1-score	support
0	0.98	1.00	0.99	51
1	1.00	0.99	0.99	82
accuracy			0.99	133
macro avg	0.99	0.99	0.99	133
weighted avg	0.99	0.99	0.99	133

- **DECISION TREE**

The classification report:

	precision	recall	f1-score	support
False	0.82	1.00	0.90	51
True	1.00	0.87	0.93	82
accuracy			0.92	133
macro avg	0.91	0.93	0.92	133
weighted avg	0.93	0.92	0.92	133

- **SUPPORT VECTOR MACHINE**

The classification report:

	precision	recall	f1-score	support
False	0.89	1.00	0.94	51
True	1.00	0.93	0.96	82
accuracy			0.95	133
macro avg	0.95	0.96	0.95	133
weighted avg	0.96	0.95	0.96	133

ROC AUC SCORE:-

- RANDOM FOREST CLASSIFIER = 1.0
- DECISION TREE CLASSIFIER = 0.9398613103778096
- SUPPORT VECTOR MACHINE CLASSIFIER = 1.0

F1 Score (weighted):-

- RANDOM FOREST CLASSIFIER = 0.9924946382275899
- DECISION TREE CLASSIFIER = 0.9183456646126856
- SUPPORT VECTOR MACHINE CLASSIFIER = 0.955283779067923

BEST MODEL:-

Decision Tree classifier Performs best