

## Data Mining & Machine Learning

### 20 Marks Component

SL No.	Tasks
1.	Regression Algorithms
2.	Regression Algorithms
3.	<b>Apriori Algorithm</b>
4.	<b>Decision Trees</b> 1. <b>ID3</b> → (extension of D3) 2. <b>C4.5</b> → (successor of ID3) 3. <b>CART</b> → (Classification And Regression Tree) 4. <b>CHAID</b> → (Chi-square automatic interaction detection Performs multi-level splits when computing classification trees)
5.	<b>KNN Algorithm</b> Use any one of the following Distance Measures 1. Euclidean Distance 2. Manhattan Distance 3. Minkowski Distance 4. Hamming Distance
6.	<b>Support Vector Machines</b>
7.	<b>Naive Bayes</b> Try to implement any one of the following variants 1. Gaussian Naive Bayes 2. Multinomial Naive Bayes 3. Bernoulli Naive Bayes
8.	<b>Clustering Algorithms</b> Try to implement any one of the following algorithm. 1. K-Means 2. K-Medoids Use the diabetic dataset: Dataset: <a href="https://www.kaggle.com/uciml/pima-indians-diabetes-database">https://www.kaggle.com/uciml/pima-indians-diabetes-database</a>
9.	<b>Clustering Algorithms</b> Try to implement any one of the following algorithm. 1. DBScan 2. Grid Based Clustering 3. Mean-Shift Clustering Use the diabetic dataset: Dataset: <a href="https://www.kaggle.com/uciml/pima-indians-diabetes-database">https://www.kaggle.com/uciml/pima-indians-diabetes-database</a>
10.	<b>Clustering Algorithms</b> Try to implement any one of the following algorithm. 1. Agglomerative Hierarchical Clustering 2. Fuzzy C Means Clustering Use the diabetic dataset: Dataset: <a href="https://www.kaggle.com/uciml/pima-indians-diabetes-database">https://www.kaggle.com/uciml/pima-indians-diabetes-database</a>