**Practical 1**

**Aim:** Implementation of supervised learning algorithms from scratch.

**Problem Definition:** Download a labelled dataset (.csv) and implement different supervised learning algorithms, such as Decision Tree, K-NN, and Naive Bayes. Understand the methodologies, implementation details, and libraries of each of the learning algorithms. Compare the performance of each algorithm for the given dataset.

**Theory:**

**Decision Tree:**

* Decision trees are a non-linear supervised learning algorithm that is used for both classification and regression tasks.
* They partition the data into subsets, where each subset corresponds to a specific value of a feature.
* The decision tree algorithm makes decisions by splitting the dataset based on features that lead to the most information gain.

**K-Nearest Neighbor (K-NN):**

* K-NN is a simple yet effective classification algorithm used for both supervised and unsupervised learning.
* It classifies data points based on their similarity to other data points in the training dataset.
* The "K" in K-NN refers to the number of nearest neighbors that influence the classification of a given data point.

**Naive Bayes:**

* Naive Bayes is a probabilistic classification algorithm based on Bayes' theorem.
* It is used for text classification and spam filtering.
* Naive Bayes makes an assumption that features are conditionally independent, which is why it is called "naive."

**Algorithms:**

**Decision Tree:**

* Import the necessary libraries.
* Load the dataset using the **importdata** function.
* Split the dataset into training and testing sets.
* Train the Decision Tree classifier using both Gini Index and Entropy as criteria.
* Make predictions using the trained classifiers.
* Calculate and display accuracy, confusion matrix, and classification report.

**K-Nearest Neighbor (K-NN):**

* Import the necessary libraries.
* Load the dataset using the **importdata** function.
* Split the dataset into training and testing sets.
* Use StandardScaler to scale the features.
* Find the best K value using cross-validation.
* Create a K-NN classifier with the best K value.
* Make predictions using the trained classifier.
* Calculate and display accuracy, confusion matrix, and classification report.

**Naive Bayes:**

* Import the necessary libraries.
* Load the dataset using the **importdata** function.
* Split the dataset into training and testing sets.
* Initialize and train a Naive Bayes classifier.
* Make predictions using the trained classifier.
* Calculate and display accuracy.

**Result Analysis:**

For each algorithm (Decision Tree, K-NN, and Naive Bayes), you need to analyze the following:

* The predicted labels for the test data.
* Accuracy score.
* Confusion matrix.
* Classification report, including precision, recall, and F1-score.

**Conclusion:**

In the conclusion, you should summarize the results and make comparisons between the three algorithms. Discuss their performance on the given dataset and any insights gained from the analysis. You can also suggest areas for further improvement or other algorithms to consider for the dataset.