K-Nearest Neighbours

**Objective:**

The objective of this assignment is to implement and evaluate the K-Nearest Neighbours algorithm for classification using the given datasets

**Dataset:**

Need to Classify the animal type

**Tasks:**

1. Analyse the data using the visualizations

2. Preprocess the data by handling missing values & Outliers, if any.

3. Split the dataset into training and testing sets (80% training, 20% testing).

4. Implement the K-Nearest Neighbours algorithm using a machine learning library like scikit-learn On training dataset

5. Choose an appropriate distance metric and value for K.

6. Evaluate the classifier's performance on the testing set using accuracy, precision, recall, and F1-score metrics.

7. Visualize the decision boundaries of the classifier.

**Interview Questions:**

1. What are the key hyperparameters in KNN?

**Ans)** The most important hyperparameter for KNN is the **number of neighbors (n\_neighbors)**.

Also hyperparameters include the number of nodes and layers in a neural network and the number of branches in a decision tree. Hyperparameters determine key features such as model architecture, learning rate, and model complexity.

1. What distance metrics can be used in KNN?

**Ans)** Minkowski, Euclidean, Manhattan, Chebyshev, Cosine, Jaccard, and Hamming distance were applied on kNN classifiers for different k values. It is observed that Cosine distance works better than the other distance metrics on star categorization.