K-Nearest Neighbours

**Objective:**

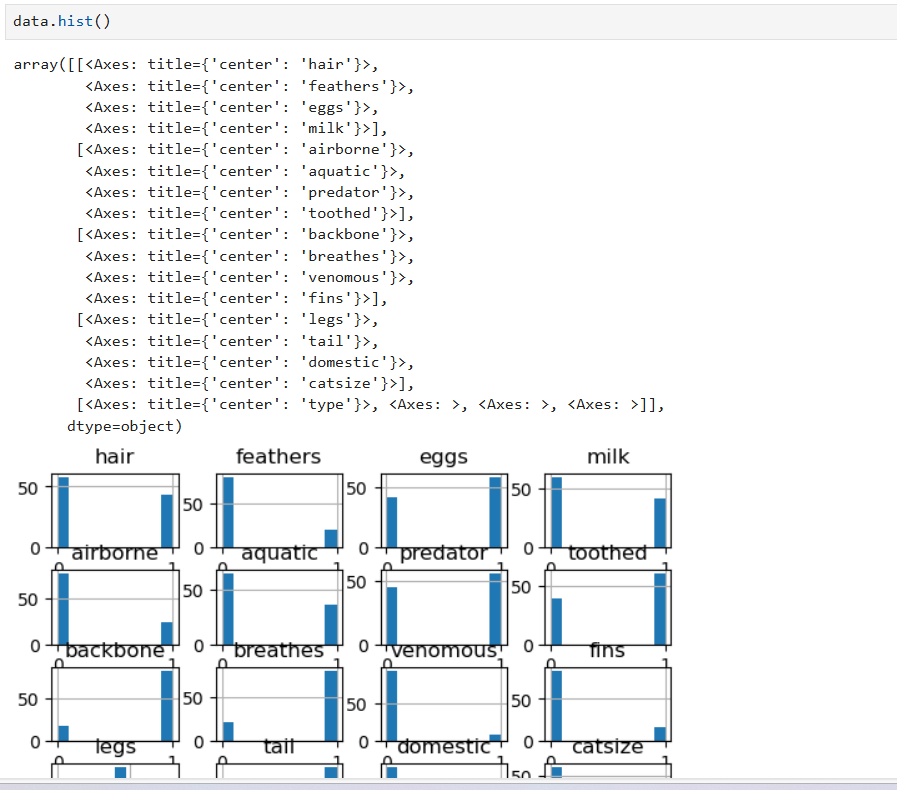
The objective of this assignment is to implement and evaluate the K-Nearest Neighbours algorithm for classification using the given datasets

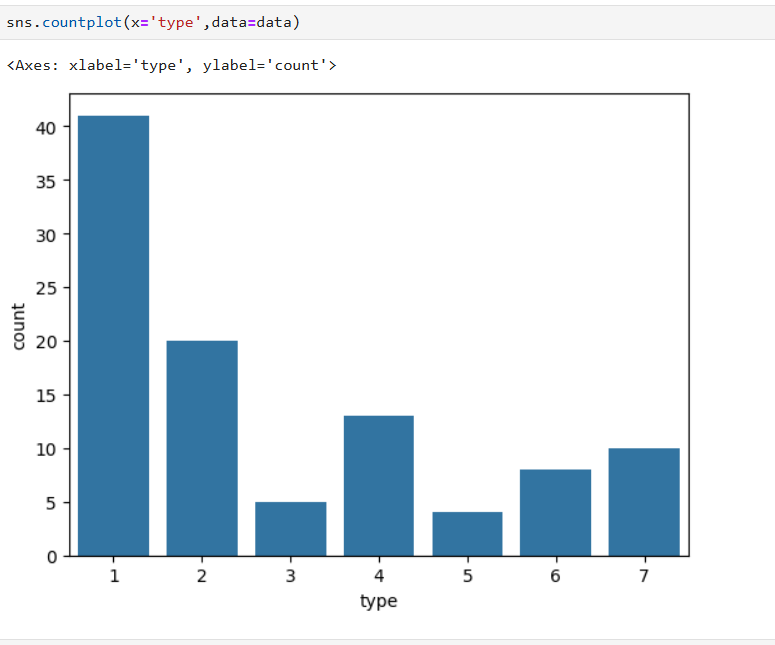
**Dataset:**

Need to Classify the animal type

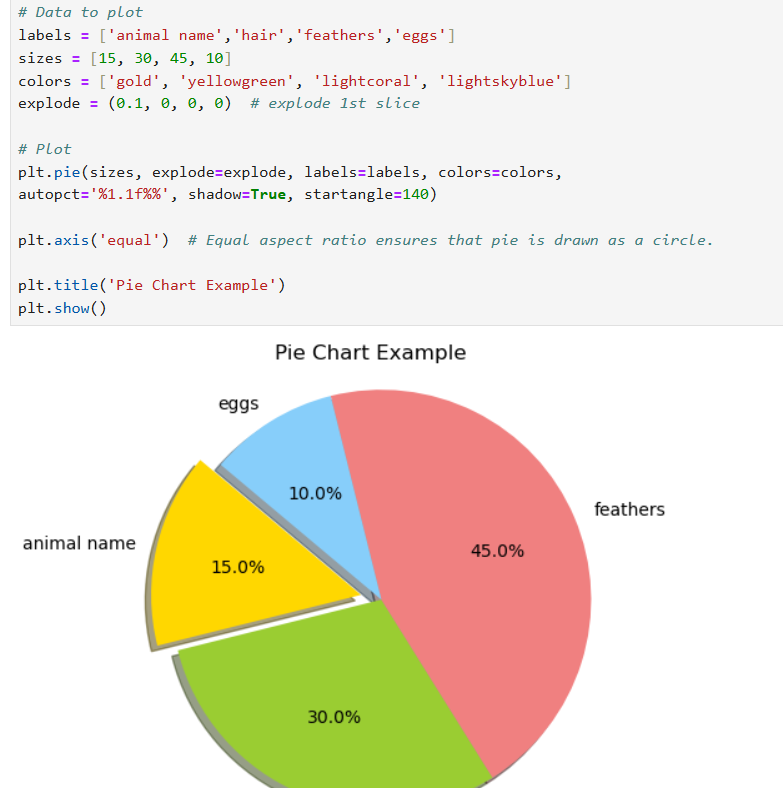
**Tasks:**

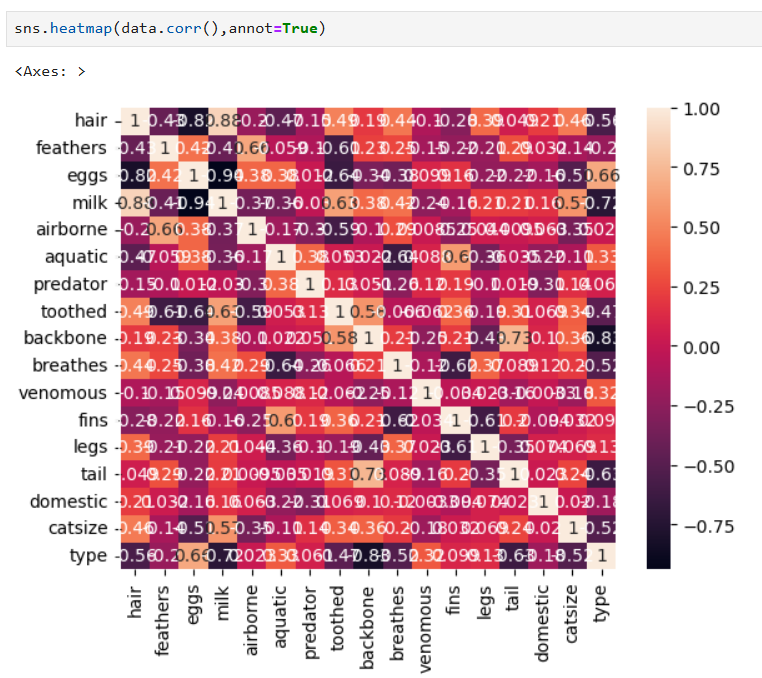
1. Analyse the data using the visualizations



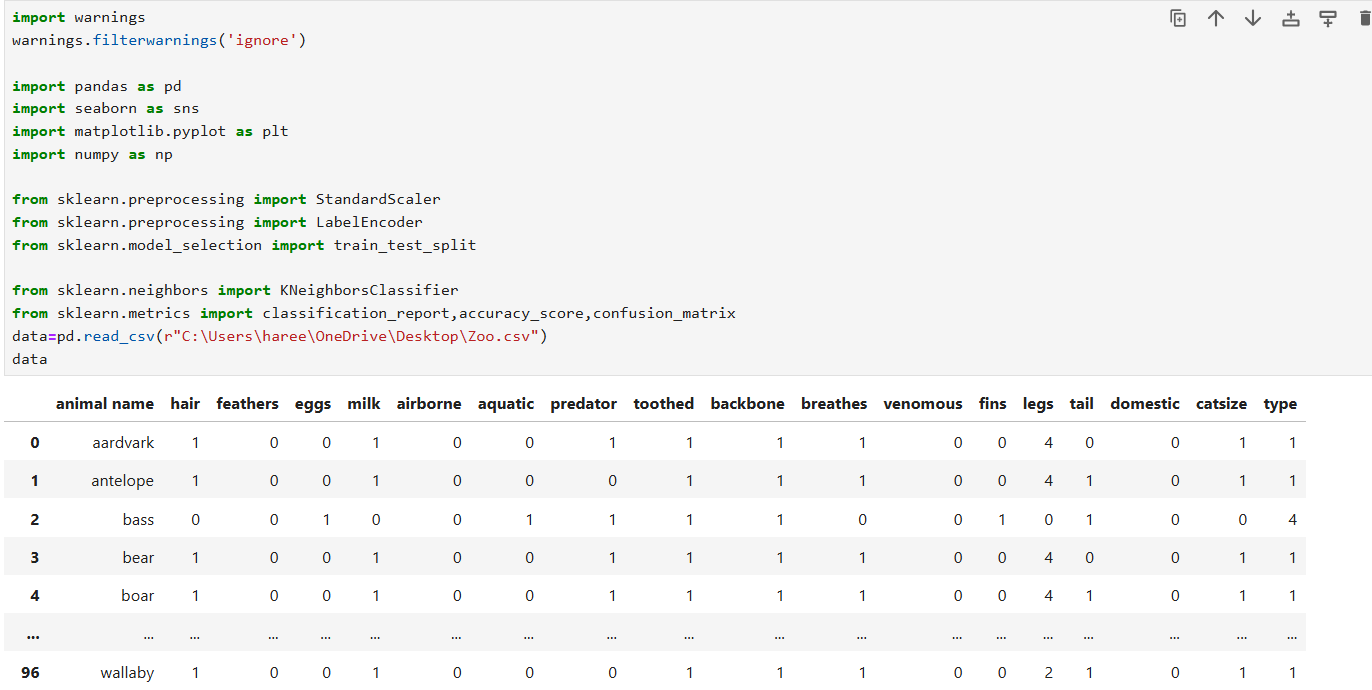


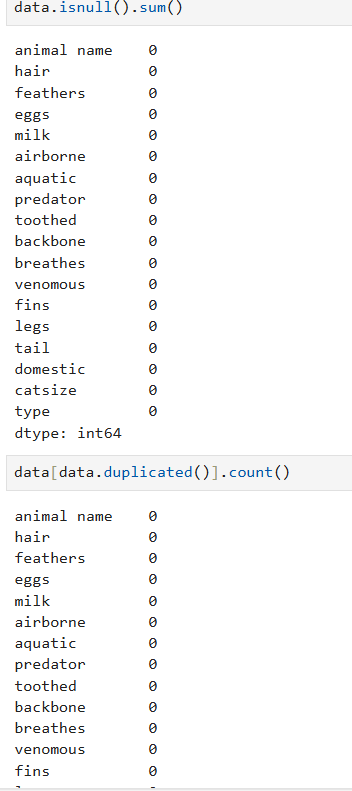




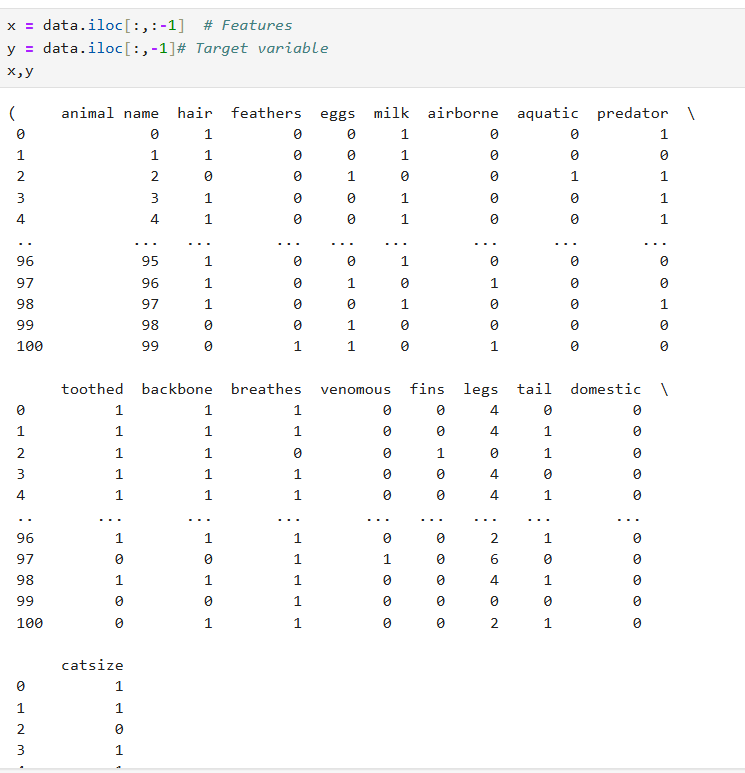


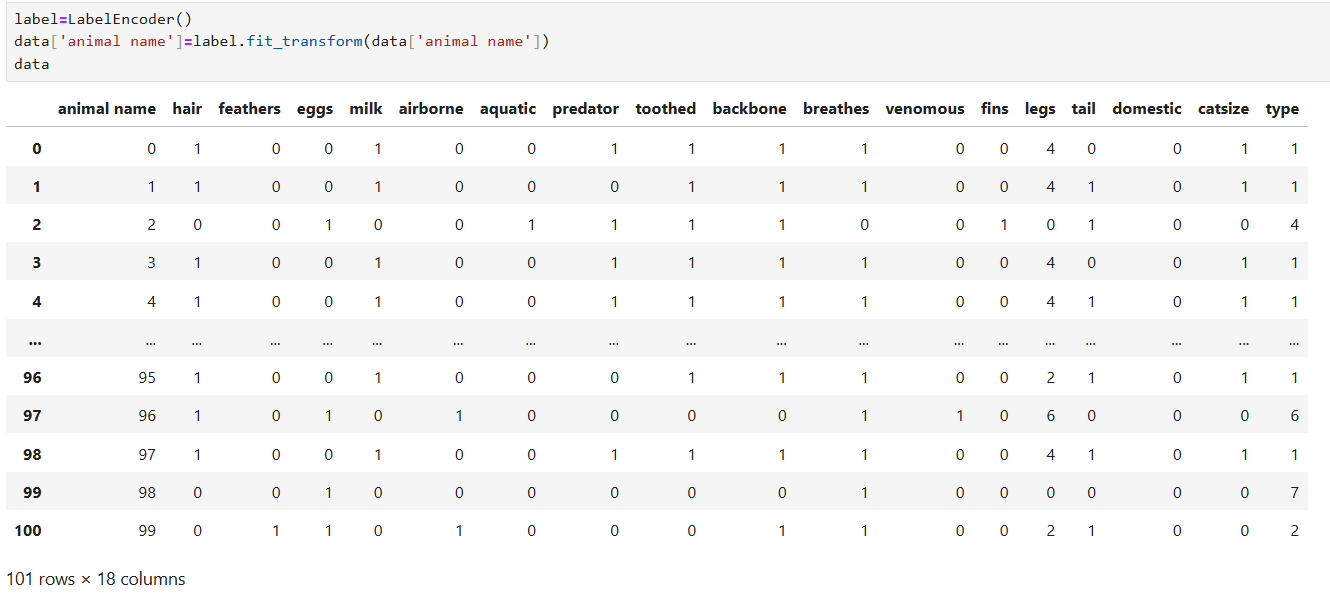
1. Preprocess the data by handling missing values & Outliers, if any.



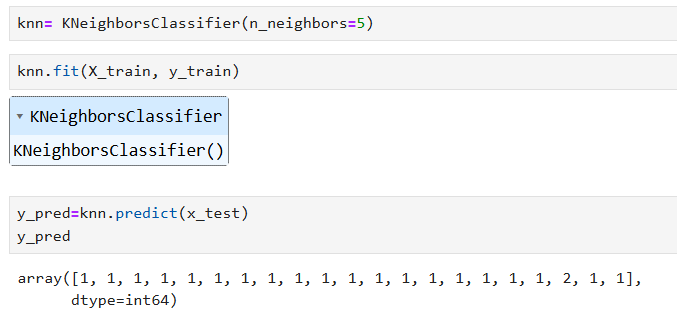


3. Split the dataset into training and testing sets (80% training, 20% testing).



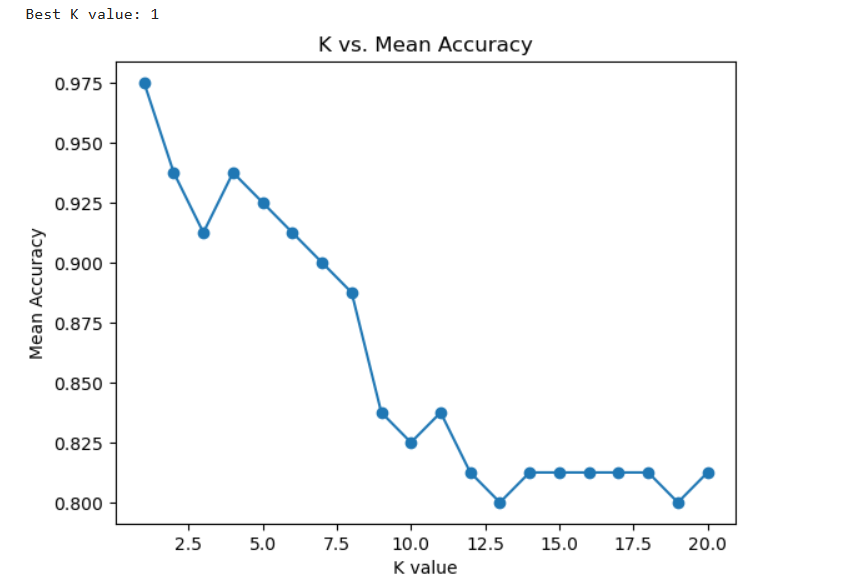


4. Implement the K-Nearest Neighbours algorithm using a machine learning library like scikit-learn On training dataset

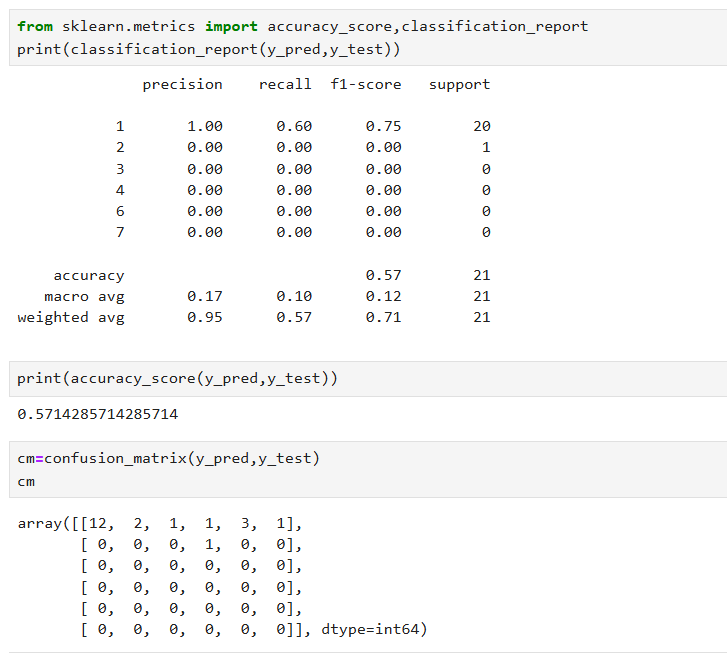


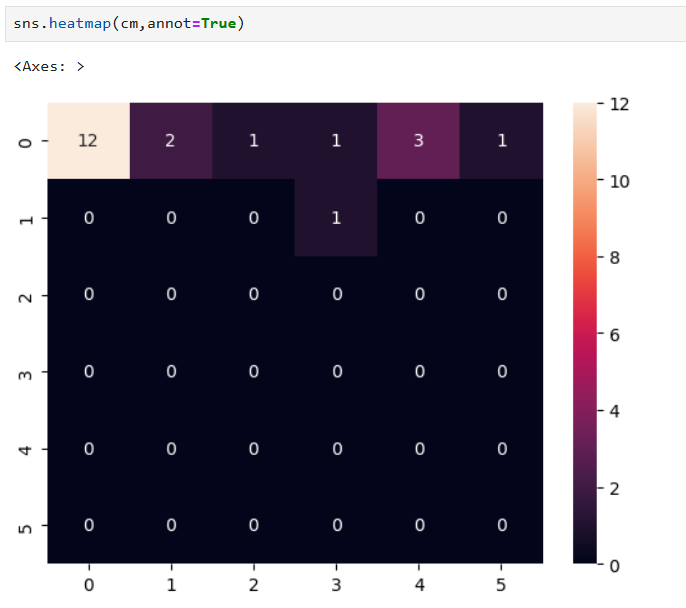
5. Choose an appropriate distance metric and value for K.



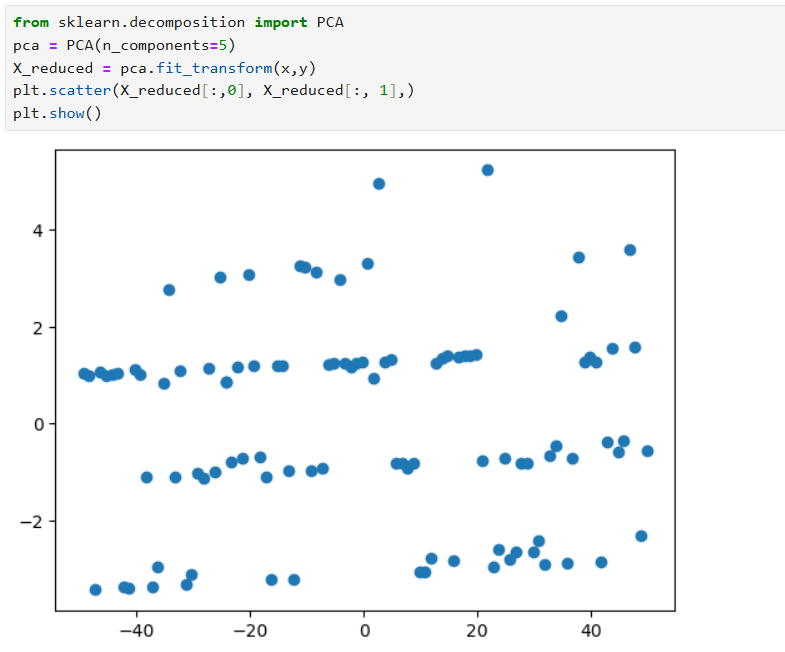


6. Evaluate the classifier's performance on the testing set using accuracy, precision, recall, and F1-score metrics.





7. Visualize the decision boundaries of the classifier.



**Interview Questions:**

1. What are the key hyperparameters in KNN?

Answer : In the K-Nearest Neighbours (KNN) algorithm, hyperparameters are parameters that are set prior to the training process and are not learned from the data. The key hyperparameters in KNN are; Number of neighbours (k),Distance metric, Weights, Algorithm, Leaf size

1. What distance metrics can be used in KNN?

Answer : In K-Nearest Neighbours (KNN), distance metrics are used to measure the similarity or dissimilarity between data points. Several distance metrics can be employed in KNN, depending on the nature of the data and the problem domain. Some common distance metrics used in KNN include: Euclidean distance, Manhattan distance (City block distance or L1 distance),Minkowski distance, Chebyshev distance, Cosine similarity