# Zoo Animal Classification

*Dataset:* https://www.kaggle.com/uciml/zoo-animal-classification?select=zoo.csv
Source: UCI Machine Learning: https://archive.ics.uci.edu/ml/datasets/Zoo

The above dataset consists of 101 animals from a zoo, with 16 representative attributes for each animal, along with the classes to which they belong.

The purpose is to classify and predict the classes of animals based on their traits, and the ultimate goal is to design a webapp analogous to Akinator, in which it asks the user a series of questions and guesses the animal the user is thinking of.

*Approach:*

Data Preprocessing:

Since the dataset is fairly neat, there will not be much data pre-processing except for the possibility of attribute reduction. One or more columns would be reduced if it is observed that the corresponding feature(s) is/are not critical to our classification.

Machine Learning Model:

We are hoping to find the class to which an animal belongs given some zoo animals with their traits. Thus, this is a case of a classification model, specifically, K-nearest neighbour.

- Propose a classification model
- Train on 50% of the dataset
- Validate on 30% of the dataset
- Test on 20% of the dataset

Evaluation Metric:

Confusion matrix is a suitable evaluation method on a classification model. We are striking for an $\geq 85\%$ accuracy with such a small dataset.

Final Conceptualization:

The final demo would be a webapp that is similar to Akinator, which interacts with users to guess a zoo animal through asking questions on the traits that we trained our model on.