Homework #1 – Due Feb. 11

- Analyze the zoo dataset from the UCI repository using the Weka Explorer.
  - For each of the attributes feathers, predators, tail, and domestic, report on the types and numbers of animals having the attribute true.
  - Remove instances whose "type" attribute is larger than or equal to 4. Use the classifier J48graft to derive the corresponding decision tree. Draw the corresponding tree.
  - Use the rules classifier PART to derive the rules on the zoo dataset. List the rules obtained.
  - Remove the "type" attribute from the dataset and run the default clustering algorithm SimpleKMeans. How many clusters do you obtain? Can you relate these clusters to the initial class values?