CDS503: Machine Learning
Academic Session: Semester 1, 2018/2019

School of Computer Sciences, USM, Penang

LAB EXERCISE (LAB 3)

Lab 3 Exercise

Look at the bank note authentication data (banknote.csv). Look at each attribute and see what type of data it has. The data can be downloaded from:

http://archive.ics.uci.edu/ml/datasets/banknote+authentication

Question 1

Explore the data and answer the following:

- What **type** of attributes of the data?
- Explore the data to check, whether is it **balance**?
- What attribute is **useful** and what is **not**? **why**?

Question 2

Experiment with the decision tree algorithm introduced in this lab to **predict** the class of the bank note authentication data (banknote.csv). Use **default** parameter configurations. Try conduct also **at least** 5-fold cross-validation (refer to your previous lab). Compare the algorithms and specify your findings. Answer the following questions:

- Do **cross-validation** effect algorithms' performance?
- Try using **1 other parameter** configurations and compare them to their **default** configuration. How much they differ?

Post your solution on Lab 03 Submission on **elearn@usm**. Make sure you to include your name and lab# on the submission post.

Format: in .ipynb

The due date is **21 October 2018 23:59**