**Implementation and evaluation of a Decision Tree and a Random Forest against the SVM with RBF kernel**

**Zhenqiang Ni** zhenqiang.ni@postgrad.manchester.ac.uk

**Jiaxing Yang** jiaxing.yang@postgrad.manchester.ac.uk

The University of Manchester

**Abstract**

Decision Tree (DT) and Random Forest (RF) are two popular machine learning algorithms for multi-class classification and regression. This paper focuses on implementing the DT and using it to then implement the RF. In addition, this project also evaluates the DT and RF models against the SVM with RBF kernel. Three datasets (‘Balloons’, ‘mushroom’ and ‘lymphography’) are applied in this project and all programming, comparisons and evaluation are completed in Python and Excel.

1. **Introduction**

generalization error

Evaluation of RF against the number of trees m

Evaluation of RF against a single tree, SVM

Estimation of Variable Importance using RF

Variable Selection using RF

Figure 1. the error rate of the RF against the number of the trees (Dataset: Lymphograhy)

Figure 2. the error rate of the RF against the number of the trees (Dataset: Balloons)