# Data Mining On Distributed Assessment System

T.T. GOH
E.E. KWAN
Massey University
t.t.goh@Massey.ac.nz

## ABSTRACT

A final year project where a data mining module and an on-line assessment system were developed. The data mining module was developed based on a decision tree algorithm. The Decision algorithm has been used widely and successfully in many areas of data mining and knowledge discovery. Examples are the location of protein coding in human DNA [1], data analysis and diagnosis for space center [2], [4], and evolutionary learning [5]. The objective of the decision tree algorithm is to fine the "pure" leaf using attributes such as entropy. The on-line education assessment system provides the necessary inputs for positive and negative examples to the data mining and knowledge discovery engine. Students' profiles, school profiles and score from the on-line assessment system will be used as the input to the mining engine. The objectives are to determine a set of rules to classify the students' learning ability and to provide necessary pre-training and feedback. It can be used for the purpose of class organization and deployment of different teaching methods to different groups.

The on-line assessment system is a three-tier architecture. It breaks an application into presentation, business logic and data access logic. Tiers are logical and may be implemented on physically separate servers. Tier one is running java applet and is mainly for data handling and complex inputs. The server engine is resided in tier two. Tier two mainly handles java applet request, invoke appropriate database, initial SQL call, and validate user and compilation of result. Tier three is where the database resides. It is accessed through RMI where a light framework is maintained.

### Keywords
Data Mining, Decision Tree, Assessment system.

## References
**Salzberg S.** "Locating Protein Coding Region in Human DNA using a Decision Tree Algorithm". Johns Hopkins University.

**Michalski R.S., Bratko I., Kubat M.** "Machine Learning and Data Mining Methods and Application". Wiley 1998.

**Michie D., Spiegelhalter D.J., Taylor C.C.** "Machine Learning, Neural and Statistical Classification". Ellis Horwood 1994.

**Buntine W.** "Tree Classification Software". NASA Ames Research Center.

**Paredis J.** "Decision Tree and Neural Nets: Complementary Representations For Learning".