

A data mining approach to construct graduates employability model in Malaysia

ABSTRACT

This study is to construct the Graduates Employability Model using classification task in data mining. To achieve it, we use data sourced from the Tracer Study, a web-based survey system from the Ministry of Higher Education, Malaysia (MOHE) for the year 2009. The classification experiment is performed using various Bayes algorithms to determine whether a graduate has been employed, remains unemployed or in an undetermined situation. The performance of Bayes algorithms are also compared against a number of tree-based algorithms. Information Gain is also used to rank the attributes and the results showed that top three attributes that have direct impact on employability are the job sector, job status and reason for not working. Results showed that J48, a variant of decision-tree algorithm performed with highest accuracy, which is 92.3% as compared to the average of 91.3% from other Bayes algorithms. This leads to the conclusion that a tree-based classifier is more suitable for the tracer data due to the information gain strategy.

Keyword: Classification; Bayes methods; Decision tree; Employability