

Student: LIN LING

Percentage: _____

QUESTION	EXCEPTIONAL 5	MERTORIOUS 4	COMPETENT 3	PASS 2	FAIL 1
Dataset description (15%)	Infers multiple pre-processing AND analysis techniques from descriptive statistics	Infers pre-processing OR analysis techniques from descriptive statistics	Provides descriptive statistics for the dataset	Correctly defines the data types; describes the original data source	Misinterprets data types and meanings
Data pre-processing (10%+10%)	Performs multiple complex primary pre-processing operations	Codes data sensibly and discusses any data removal or manipulation	Performs obvious coding (Yes/No) OR justifies data manipulation	Describes the data used but does not perform any significant pre-processing	Performs no primary pre-processing
	Performs multiple secondary pre-processing operations	Recodes data to allow further analyses AND adds new calculated fields	Recodes data to allow further analyses OR adds new calculated fields	Performs appropriate secondary pre-processing	Performs no secondary pre-processing
Data analysis - data exploration (15%+10%)	Uses multiple methods other than decision trees	Uses multiple types of algorithm AND multiple options with each of these.	Uses multiple types of algorithm appropriately and shows a logical progression within an analysis	Uses the same type of algorithm in each analysis	Incorrect, misleading or missing
	Full and detailed discussion of the selected model				
	Fully tests each model in each analysis	Uses MVT methodology correctly	Includes prediction and comparison of algorithms	Only creates models on the total dataset	Incorrect, misleading or missing
Reporting of results (20%)	Four complete analyses with discussion of the implications of the results	Three complete analyses with discussion of the implications of the results	Two complete analyses with discussion of the implications of the results	One complete analysis with discussion of the implications of the results	Incorrect, misleading or missing
Reflection (10%)	Full and detailed reflection with discussion of how the analysis might be repeated and extended	Discussion of how the research might be repeated OR extended.	Discussion of satisfying analyses and problems encountered.	List of skills learned.	Incorrect, misleading or missing
Report writing (10%)	Reads like a technical review article	Good use of signposting	Clear sectioning of content	Basic use of introduction, body and conclusion; correctly referenced	No clear structure; incorrectly referenced

Always include a cover sheet with your ~~own~~ name, the name of the workshop and the name of the assignment.

Clear description of the attributes

Good handling of two data, and good interpretation of the visualisation.

Good idea to state the predicted outcome.

It would be better to show and discuss the confusion matrix as well as the percentage correct. This allows you to identify where the problems are. There's no need to use cross-validation when you have a large dataset.

You need to discuss the results more - you have numbers, but do those show any what sort of weather leads to accidents?