**Appendix D Self Assessment Rubric**



**Activity 1 Business Problem and EDA Feedback**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| ***Activity 1*** | **F** | **D** | **C** | **B** | **A** |
| ***Business Problem***  ***Data mining problem***  ***Use of Exploratory Data Analysis in getting to know data***  ***Workflow diagram***  ***Methodology*** | none | Incorrect assignment of target variable  No methodology presented  Attempted made at using exploratory data analysis but conclusions are irrelevant  Little or weak discussion. | Attempt made at identifying business problem and linking it to a data mining problem. Relevant Target variable identified.  No methodology presented  Or missing workflow diagram  Attempt made at using exploratory data analysis  No initial assessment of data provided  Some minor mistakes | Relevant Business Problem identified.  Relevant corresponding target variable identified  Business problem transferred to an appropriate data mining problem with correct justification  Methodology and workflow diagram presented and if required adapted to the business problem.  Exploratory data analysis provided but no justification presented for transformations or enriching the signal in the data | Relevant Business Problem identified. Relevant corresponding target variable identified  Business problem transferred to an appropriate data mining problem with correct justification  Methodology and workflow diagram presented and justified  Exploratory Data analysis used to identify limitations with data.  A relevant assessment of data quality provided |

**Activity 2 Regression Analysis Self Assessment Feedback**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| ***Activity 2*** | **F** | **D** | **C** | **B** | **A** |
| ***Interpretation of model results after correctly applying the Impute node and any necessary transformation-s***  ***Technical detail and the interpretation of the Results of the Full Regression model and the interpretation of the Results of the various selection methods used in Regression modelling*** | none | No explanation of use of replacement or transformation node.  Inappropriate use of transformations  No regression equation presented or incorrect interpretation of parameter estimates  No use of selection methods, if used interpretation of the results not discussed | Some models obtained using Replacement node with some explanation.  Full Regression model obtained with some explanation. Attempt made to interpret the results a regression equation presented  No updated workflow diagram provided | Appropriate models obtained using Replacement node and valid transformation node with reasonable justification.    Reasonable attempt to interpret results A valid regression equation presented with statistically significant parameters identified | Discussion on appropriate solutions to transform the data to enrich the signal discussed comprehensive exploration of transformations with a concise review of the most important transformations with supporting valid justification.  Correct, comprehensive technical detail and excellent interpretation of the results from selection methods. |

**Activity 3 Decision Tree Analysis Self Assessment Feedback**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| ***Activity 1*** | **F** | **D** | **C** | **B** | **A** |
| ***Technical detail and the interpretation of the Results of the Decision Tree model*** | none | Decision Tree obtained no interpretation provided.  Little or weak analysis  No exploration away from default parameters or manual tree explored. | Decision Tree model obtained with some explanation. Reasonable attempt made to interpret the results. Import ant target paths described . Very little exploration beyond the SAS default parameters.  No updated workflow diagram provided | Sensible model obtained exploration of decision tree node parameters resulting in an improved model. with reasonable explanation. Contrasted with a module produced by hand . All target paths correctly identified with emphasis on those rules which have the most impact in predicting target variable. | In addition to grade B discussion and evidence provide to indicate whether models have been overfitted and limitations discussed. Justification provided for the best decision tree model |

**Activity 4 Neural Network Analysis Self Assessment Feedback**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| ***Activity 1*** | **F** | **D** | **C** | **B** | **A** |
| ***Technical detail and the interpretation of the of the Neural Network model*** | none | Incorrect presentation of neural network architecture  Little or weak analysis  No assessment of important weights. | Neural Network model obtained with some explanation. Attempt made to interpret the results with a description of the neural network architecture. No exploration of SAS defaults for this node.  No updated workflow diagram provided | Good model with good explanation. Correct technical detail, correct description of neural network architecture, critical appropriate exploration of neural network parameters identification of weights which have the most impact. | Same as for Grade B analysis and justification provided to indicate if models or models are overfitted with justification provided for selection of the best neural network model. |