	D	C	B	A
Data Cleaning, Analysis and Feature Selection **0 – 2: Not very well done [Grades D or C]** **3 – 4: OK [Grades B or A]**	Missing data removed.	Missing data imputed (only applicable if less than 30% of instances are missing)	Feature selection done and justification provided in notebook	Imputation based on interacting features and/or distribution of data within feature
Model(s) / Technique(s) / Ensemble(s) **0 – 2: Not very well done [Grades D or C]** **3 – 4: OK [Grades B or A]**	Only one model tried. No explanation why it was chosen. Student has not shown evidence of putting in any thought / effort into model selection	Two Models / Techniques listed with no explanation of why these were chosen. Models work but results not compared against each other	At least one ensemble consisting of two or more models tried. Results of each model compared against each other.	Two or more ensembles tried. Graphs showing how results differ by changing composition of ensembles or voting methods shown.
Hyperparameters Tried **0 – 2: Not very well done [Grades D or C]** **3 – 4: OK [Grades B or A]**	No hyperparameters tried. Default values chosen	Two hyperparameters chosen for per model, but no analysis of which hyperparameter works better	Multiple hyperparameters chosen. Results of differences in results compared	Differences of results due to hyperparameters analysed with graphs.
Training, Validation, Testing **0 – 2: Not very well done [Grades D or C]** **3 – 4: OK [Grades B or A]**	No mention of difference between training and testing split of data. No demonstration of how split was done.	One training /validation/testing split. No other values tested	Multiple training/validation/testing splits experimented, and results shown, but no analysis of how splits affects results	Analysis of how training/validation/testing splits affect results.
Evaluation and Conclusion **0 – 2: Not very well done [Grades D or C]** **3 – 4: OK [Grades B or A]**	Simple accuracy tested; no other metric considered	Average class accuracy for multiple classes shown	Precision, recall and F1 metric (or SSE / MSE / RMSE / MAE, depending on what the student has decided to do) calculated for one model	Evaluation metrics for all models compared and best model(s) identified