	Criteria	High distinction (100% to 80%)	Distinction (70% to 79% of available mark)	Credit (60% to 69% of available mark)	Pass (50% to 59% of available mark)	Fail (<50% of available mark)
	Implementation					
Task 1 Implementation [4 marks]	**Basic Text Pre-processing**	All required basic text pre-processing steps are performed correctly without any error. Included all required output files. Meets all filename and output format requirements.	All required basic text pre-processing steps are performed correctly, with one or two minor errors. Included all required output files. Meets all filename and output format requirements.	Most required basic text pre-processing steps are performed correctly, with one or two minor errors. One or two steps are missing, or not done properly. Included all required output files. Meets most filename and output format requirements.	Some required basic text pre-processing steps are performed correctly, with some minor errors. Some steps are missing, or not done properly. Included most required output. Some output files do not meet the naming or format requirements.	Few to none text pre-processing steps are performed, or many of them have major errors. Included minimum or no required output file. Many output files do not meet the format requirements.
Task 2 Implementation [7 marks]	**Generating Feature Representations**	All required document feature representations are generated properly. The methods used are justifiable, proper and effective. The output count vector meets all the format requirements.	All required document feature representations are generated properly. The methods used are mostly correct and justifiable. The output count vector meets all the format requirements.	Most required document feature representations are generated. Most methods used are correct and justifiable, some are questionable and/or without a not clear purpose. The output count vector meets most of the format requirements.	At least half of the required document feature representations are generated. Some methods used are correct and justifiable, many are questionable and/or without a not clear purpose. The output count vector meets most of the format requirements.	Few to none of the required document feature representations are generated. The methods used are not correct nor effective. The output count vector fails to meet most of the format requirements.

Task 3 Implementation [7 marks]	**Job Advertisement Classification**	The nature of all tasks has been properly investigated and discussed before building machine learning models. All required models are built properly. All research questions are answered correctly supported by comprehensive experiments and robust evaluations.	The nature of each task has been properly investigated and discussed before building machine learning models. Most required models are built properly. Most research questions are answered correctly supported by comprehensive experiments and robust evaluations.	The nature of most tasks has been properly investigated and discussed before building machine learning models. Many required models are built properly. Most research questions are answered correctly supported by complete experiments and evaluations.	Some required models are built properly. Many models are built without initial investigations on the nature of the tasks. Some research questions are answered correctly supported by experiments and evaluations. Some evaluations are incomplete or without robust comparisons.	Few to none models are built correctly. There is no initial investigation on the nature of the tasks. Few to none research questions are answered, or most of the research questions are answered incorrectly or without experiments or evaluations.
	Notebook Presentation					
Notebook Presentation	**Notebook Presentation| Code Commenting**	Commenting provides clear explanations of how the code is intended to work. It also provides clear instructions and insights to why the code has been written as it has. Commenting is thorough, and concise. Comments provided would be helpful to other data scientists.	Commenting provides useful insights into how the code is intended to work. It also provides good instructions as to why the code has been written as it has. Commenting has been provided in most cases and is generally concise. Most comments provided would be helpful to other data scientists.	Commenting provides some explanation as to how the code is intended to work. It also provides adequate instructions as to why the code has been written as it has. Commenting has been provided in most cases. Most comments provided would be helpful to other data scientists with some corrections or additions made.	Commenting provides some explanation as to how the code is intended to work, but is at times unclear. Some instruction is provided, but should have provided more details. Commenting has been provided in some of the expected cases. Comments provided have some usefulness to other data scientists, but require corrections or additions to be made.	Code commenting does not provide a clear explanation of the code. There is a lack of instructions as to why the code has been written as it has. Commenting is minimal OR not provided at all.

	Notebook Presentation | Notebook Content	The jupyter notebook provides comprehensive justifications and discussions on the methods/approaches chosen, results, analysis and findings. Effective use of markdown, with clear sectioning, highlight, styling, etc. The jupyter notebook is absolutely clear, concise, exceptionally well organized, has clear logic flow and very easy to follow.	The jupyter notebook provides most of the justifications and discussions on the methods/approaches chosen, results, analysis and findings. Effective use of markdown, with clear sectioning, highlight, styling, etc. The jupyter notebook is mostly clear and concise. It is in general well organized, shows some logic flow, and fairly easy to read.	The jupyter notebook provides many justifications and discussions on the methods/approaches chosen, results, analysis and findings. The jupyter notebook contains some irrelevant information. It is mis-leading in some places. It is readable by someone who knows what it is supposed to be doing. However, it requires improvement on the presented logic flow.	The jupyter notebook provides some justifications and discussion on the methods/approaches chosen, results, analysis and findings. The jupyter notebook contains lots of irrelevant information. It is mis-leading in lots of places. It is not well organised and there is no clear logic flow.	The jupyter notebook provides minimal justification and discussion on the methods/approaches chosen, results, analysis and findings. The jupyter notebook contains lots of irrelevant information. It is poorly organized, very messy, and is not readable.
	Note: - Rubric for Notebook Presentation (including code commenting and notebook content) is common across Task 1, 2 and 3. - Students can refer to the activities in modules as examples for the level of details that they should include in their jupyter notebook.					