

Student ID

Overall mark #NAME?
Mark on the categorical scale

Note If you submitted late and do not have a valid ECF (Extenuating Circumstances Form), your mark will be capped at 50.
The marks are subject to moderation and approval by the Module Assessment Board and the Board of Examiners.

Marking scheme	Marks	under 30%				30%			40%			50%			60%			70-80%				85% and above					
		0	15	22	28	32	35	38	42	45	48	52	55	58	62	65	68	72	75	78	82	85	88	95	100		
1. [30%] Preprocess the textual data to remove potential noise. Use, as a minimum, the following preprocessing methods: 1) removing punctuation, 2) removing numbers, 3) removing stop words, 4) changing the text to lower/upper case and 5) lemmatising. Describe in detail, with at least 3 examples, the changes to the text for each preprocessing method applied.																											
65% - Analysis of the results	19.5	0.00	2.93	4.29	5.46	6.24	6.83	7.41	8.19	8.78	9.36	10.14	10.73	11.31	12.09	12.68	13.26	14.04	14.63	15.21	15.99	16.58	17.16	18.53	19.50		
		no results displayed	results displayed for 1 preprocessing method		results displayed for 2 preprocessing methods	results displayed for 3 preprocessing methods	results displayed for 4 preprocessing methods	results displayed for 5 preprocessing methods																			
		no analysis	correct analysis for 1 preprocessing method + incorrect/ missing for 4 methods	correct analysis for 2 preprocessing methods + incorrect/ missing for 3 methods	correct analysis for 3 preprocessing methods + incorrect/ missing for 2 methods	correct analysis for 4 preprocessing methods + incorrect/ missing for 1 method	correct analysis for 5 preprocessing methods																				
20% - Use of tables and figures	6	0.00	0.90	1.32	1.68	1.92	2.10	2.28	2.52	2.70	2.88	3.12	3.30	3.48	3.72	3.90	4.08	4.32	4.50	4.68	4.92	5.10	5.28	5.70	6.00		
		no tables/ figures	few and mostly unclear/irrelevant tables/ figures	mostly unclear tables/ figures	mostly clear tables/figures AND some unclear tables/figures	mostly clear tables/figures AND very few unclear tables/figures	clear tables/figures																				
15% - Structure and presentation	4.5	0.00	0.68	0.99	1.26	1.44	1.58	1.71	1.89	2.03	2.16	2.34	2.48	2.61	2.79	2.93	3.06	3.24	3.38	3.51	3.69	3.83	3.96	4.28	4.50		
		task is missing from report	seriously deficient in organisation, style and clarity	deficient in organisation, style and clarity	acceptable organisation, style and clarity	average organisation, style and clarity	good organisation, style and clarity	very good organisation, style and clarity																			
2. [30%] Using the bag-of-words/terms representation, perform classification using 3 algorithms; present and discuss the results; compare the results of the 3 algorithms.																											
65% - Analysis of the results	19.5	0.00	2.93	4.29	5.46	6.24	6.83	7.41	8.19	8.78	9.36	10.14	10.73	11.31	12.09	12.68	13.26	14.04	14.63	15.21	15.99	16.58	17.16	18.53	19.50		

65% - Analysis of the results	19.5	0.00	2.93 4.29 5.46 10% of topics displayed	6.24 6.83 7.41 30% of topics displayed	8.19 8.78 9.36 50% of topics displayed	10.14 10.73 11.31 70% of topics displayed	12.09 12.68 13.26 all topics displayed	14.04 14.63 15.21 15.99 additional results beyond what was requested	16.58 17.16 18.53 19.50 substantially more results beyond what was requested			
			no analysis	correct analysis for 10% topics + incorrect/ missing analysis for the rest of the topics	correct analysis for 30% of topics + incorrect/ missing analysis for the rest of the topics	correct analysis for 50% topics + incorrect/ missing analysis for the rest of the topics	correct analysis for 70% of topics + incorrect/ missing analysis for the rest of the topics	correct analysis for all topics	additional insights beyond basic interpretation			
20% - Use of tables and figures	6	0.00	0.90 1.32 1.68 few and mostly unclear/irrelevant tables/figures	1.92 2.10 2.28 mostly unclear tables/figures	2.52 2.70 2.88 mostly clear tables/figures AND some unclear tables/figures	3.12 3.30 3.48 mostly clear tables/figures AND very few unclear tables/figures	3.72 3.90 4.08 clear tables/figures	4.32 4.50 4.68 4.92 tables/figures displaying additional insights	5.10 5.28 5.70 6.00 tables/figures displaying substantially more insights			
15% - Structure and presentation	4.5	0.00	0.68 0.99 1.26 seriously deficient in organisation, style and clarity	1.44 1.58 1.71 deficient in organisation, style and clarity	1.89 2.03 2.16 acceptable organisation, style and clarity	2.34 2.48 2.61 average organisation, style and clarity	2.79 2.93 3.06 good organisation, style and clarity	3.24 3.38 3.51 3.69 very good organisation, style and clarity	3.83 3.96 4.28 4.50 excellent organisation, style and clarity			
Adjustment			-5	-4	-3	-2	-1	1	2	3	4	5

Comments