


ASSIGNMENT ASSESSMENT

Criterion	5 or lower	6	7	8	9 or 10
Research question (RQ) Weight 15%	Unclear, confusing, or simplistic RQ. Mechanic or superficial RQ. RQ not motivated. RQ does not leverage potential of data or NLP (for example a simple summary of one or two variables is enough to answer the question)	Structured research question but not always adequately written nor clear. Lacks relevance or realism. Mechanic question.	Structured research question.	Good research question: well structured, right level of detail. Realistic and reasonably relevant, accurate, and to the point.	Excellent research question. Clear, relevant and allows to explore and extract the most from the data that is available. Right level of detail.
Data Weight 15%	The assignment does not explain nor motivates the data aggregation strategy nor how it was structured or organized. The connection between the RQ and data is not clear or inadequate. The connection between the data and analysis is not clear. Used data as is.	The data choices address sufficiently the RQ. It does not reflect all material covered, It includes considerable simplifications or shortcuts. Minimum level of adequacy for study load.	The data choices clearly address the RQ. Good ability to apply the knowledge covered, with a few oversights.	Well-considered and well-explained answer. Clear evidence of very good knowledge and understanding of the material covered, without being exceptional. Good use of the given data	The data addresses fully the research question and is consistent with the analysis. Creative decisions and choices regarding variables and dataset structure and aggregation providing room for insights.
Classifiers - content Weight 10%	Insufficient ability to apply or explain the knowledge covered regarding the NBC. Content likelihood computation not explained/not understandable. Major flaws in variable construction or explanation of constructed variables. Lacks motivation for decisions.	Sufficient ability to apply the knowledge covered. Despite several flaws in the assignment, the outcome is satisfactory. It includes considerable simplifications or shortcuts.	Good ability to apply the knowledge covered, with a few oversights.	Very good ability to apply the knowledge covered in the course.	Excellent calculation, explanation and interpretation of the likelihoods and NBC including variable construction, with decisions clearly explained and motivated.
Classifiers – sentiment Weight 10%	Insufficient ability to apply the knowledge covered regarding the NBC. Sentiment likelihoods not explained/not understandable. Major flaws in variable construction or explanation of constructed variables. Lacks motivation for decisions.	All words have the same likelihoods or calculation is unclear. It includes considerable simplifications or shortcuts. Despite several flaws in the assignment, the outcome is satisfactory.	Good ability to apply the knowledge covered, with a few oversights. Very few simplifications or shortcuts	Very good ability to apply the knowledge covered in the course.	Excellent calculation, explanation and interpretation of the sentiment likelihoods chosen, including motivation, calculation of the word likelihoods for sentiment and interpretation of the likelihoods and obtained posteriors.
Model evaluation – confusion matrix metrics Weight 10%	Insufficient ability to apply the knowledge covered regarding the model evaluation/confusions matrix. Calculations not explained/not understandable. Major flaws. Lacks interpretation. Prediction/validation splits overlap or confusing	it includes considerable simplifications or shortcuts. Despite several flaws in the assignment, the outcome is satisfactory. Used data splits	Good ability to apply the knowledge covered, with a few oversights. Used data splits.	Very good ability to apply the knowledge covered in the course. Good interpretation and reflection of the performance obtained and its reasons Data splits properly used.	Excellent evaluation of the performance of the classifier(s) including calculation, presentation, and interpretation of the likelihoods and obtained posteriors. Proper splits. Content NBC algorithm performed among the top of the class
Forecast tasks Weight 15%	 major flaws in them, or not possible to understand how they were done. Lacks interpretation.	it includes considerable simplifications or shortcuts. Minimum level of adequacy for study load.	Good ability to apply the knowledge covered, with a few oversights.	Very good ability to apply the knowledge covered in the course (e.g., via constructed variables). Good interpretation and reflection of the results and process.	Excellent ability to produce forecast using the constructed variables, taking into account the structure of the data (e.g., time, level of aggregation, additional variables, etc), very insightful interpretation and reflection.

Analysis Weight 20%	RQ not answered. Analysis does not align with the RQ. Mechanic analysis. Lacks interpretation. Simplistic, sometimes due to the weak RQ (e.g., does not require sentiment per topic or richer analyses/work)	it includes considerable simplifications or shortcuts. Minimum level of adequacy for study load. Despite several flaws in the assignment, the outcome is satisfactory.	Good ability to apply the knowledge covered, with a few oversights. Very few simplifications or shortcuts. It shows good knowledge of the material covered, awareness and own reflection of the important aspects of the material covered.	Good ability to apply the knowledge covered. It shows good knowledge of the material covered, awareness and own reflection of the important aspects of the material covered.	Excellent ability to apply the knowledge covered in the program given settings/data.
Written communication Weight 5%	Unstructured text. Fails to convey the key message of the thesis and/or to address questions.	The text is somewhat unstructured and unclear. More polishing work is needed.	Overall well written, with occasional typos, or inaccuracies. The writing style is mechanical.	Structured text. The text is clear and concise, but here and there more (or less) details could improve the readability. Tables and Figures are self-explanatory and timely introduced in the text. The writing style is a bit mechanical at times.	Structured, coherent and polished text. Excellent writing style. The text is accurate, clear and concise, with the right level of detail. Tables and Figures are self-explanatory and timely introduced in the text.