DSO599 – Text Analytics & Natural Language Processing Rubric

Section	Criterion	Below Expectations	Meets Expectations	Exceeds Expectations	Score
Final Presentation & Group Collaboration (Week 8) **40 points**	**Content is delivered in a logical, ordered manner that tells a coherent story**	*Content starts, transitions, and ends abruptly, with no apparent introduction or context, and not all members are present. (0)*	*Audience is confused by certain portions of presentation / Q&A (5)*	*Discussion points and recommendations are properly contextualized and build off prior concepts. (10)*	10
	Content and delivery is suitable for business decision makers, but contains content	*Presentation has multiple slides / occurrences of technical concepts / jargon introduced with no context and is meant more for a software engineer than a business manager. (0)*	*Certain slides / concepts presented may require reworking or rephrasing to generate managerial buy-in. (5)*	*Content presented is accessible, logical, and clear for business managers. (10)*	10
	Group demonstrates effective collaboration and teamwork	*360 reviews show lack of collaboration, communication, and opinions differ drastically between group members (0)*	*360 reviews show most group members collaborating effectively (5)*	*360 reviews show group documents individual contributions, assigns work fairly and equally (10)*	10
	Presentation follows 8 min time limit, at least 5 meaningful questions are asked by team members	*At least two of the three conditions are not met. (0)*	*One of the three conditions is not met. (5)*	*All three conditions are met. (10)*	10
Business Domain Solutions **30 points**	**Recommendations address a concrete business / operational problem**	*Recommendations do not address any real-life relevant issues for clients (0)*	*Recommendation addresses a concrete business problem, but implementation is unclear (7.5)*	*Recommendations provide a concrete solution to a business problem with implementation roadmap clearly articulated (15)*	20
	The Return on Investment of	*At least two of the following are missing /*	*Costs, risks, benefits, and implementation roadmap /*	*Quantified, logical projections for costs, risks,*	20

University of Southern California, Marshall School of Business

	recommendation is **appropriately quantified and communicated**	*unclear: costs, risks, benefits, and implementation roadmap / timeline. (0)*	*timeline are provided but may not be logically sound or realistic. (7.5)*	*benefits, and roadmap / timeframe for implementation are provided. (15)*	
Technical Implementation **30 points**	**Algorithms are selected/implemented properly, and produce strong results (model evaluation metrics are above baseline)**	*Wrong algorithm is selected AND implemented with substantial issues that compromise accuracy. Instructor is unable to reproduce team results. (0)*	*Correct algorithm is selected, but implementation contains small mistakes / does not handle edge cases, or simply reuses a 3rd party model w/ no modifications. (10)*	*Algorithm is correctly selected/implemented, and is modified to suit the particular use case. (20)*	20
	Solution is scalable for 5MB, 50MB, and 50GB text data	*No implementation is provided for handling larger-size data and is only performant on small flat file. (0)*	*Source code / guidelines are provided to handle larger text corpuses but contain flaws in implementation. (2.5)*	*Source code is uploaded to Github, and any data files are stored in S3. (5)*	5
	Solutions are well-documented, adhere to data science / engineering best practices	*No comments or documentation is provided. A data scientist / analyst continuing the project 6 months from now must invest significant time understanding what is happening. (0)*	*Source code contains inconsistent comments / documentation. (2.5)*	*A data scientist / analyst would be able to jump right into the project and continue working. (5)*	5
				TOTAL	100