

CIS*4500 (Fall 2019) Marking Scheme for the Project		Marks	Comments
Data Analysis (15)	Document lengths in sentences (5)		
	Document lengths in tokens (5)		
	Avg sentence lengths for collection and each document (5)		
Feature selection (10)	At least two different methods (2 x 5)		
ML Classifiers (15)	At least three ML methods (3 x 5)		
Model tuning/ optimization (24)	Proper partition of data for cross-validation (6)		
	3 ML x 2 FS x 6 Size = 36 configurations (18)		
Results and analysis (16)	Impacts of feature sizes, selection methods, ML classifiers, and their interactions (4 x 4)		
Project report (20)	Introduction (3)		
	Technique descriptions (3)		
	Implementation highlights (3)		
	Data, results, and analysis (5)		
	Conclusions and future work (2)		
	Build/test guide (2)		
	Title and references (2)		
Total (100)			