Samuel Miller

Text Processing Neural Networks

Homework 2

System	ROUGE-1			ROUGE-2			ROUGE-L		
	Precision	Recall	F-Score	Precision	Recall	F-Score	Precision	Recall	F-Score
Centroid	18.27	23.99	20.69	1.15	1.53	1.31	16.37	21.51	18.54
DPP	21.427	22.50	21.94	1.47	1.55	1.50	17.86	18.76	18.29
ICSIsumm	21.13	20.02	20.56	1.37	1.30	1.34	19.21	18.21	18.69
LexRank	18.77	23.77	20.95	1.23	1.56	1.37	17.04	21.58	19.02
Submodular	21.29	21.98	21.62	1.43	1.48	1.46	19.25	19.87	19.55

This experiment was conducted using the pyrouge library, a Python package that simply wraps the ROUGE toolkit which was originally written in PERL. To set up the development environment for this experiment, pyrouge and ROUGE-1.5.5 were installed along with their dependencies including Perl. Then, pyrouge was configured by setting its path to the directory containing the ROUGE-1.5.5 installation. Preprocessing for this experiment simply involved initializing a Rouge155 toolkit object and defining filename patterns and paths that point to the system and reference summaries. The function Rouge155.convert_and_evalute() handles the conversion of summaries to Rouge format and performs evaluation in a single step. This process was performed in a for loop for each summarization technique: Centroid, DPP, ICSISumm, LexRank, and Submodular. The results were stored in a dictionary that maps the summarization technique to the precision, recall, and f-score evaluations for Rouge-1, Rouge-2, and Rouge-L. In terms of runtime, the techniques ordered from best to worst execution time are ICSIsumm with 306.72s, DPP with 311.12s, Submodular with 321.63s, LexRank with 368.19s, and Centroid with 376.75s.