
Trainable Decoding of Sets of Sequences for Neural Sequence Models

Ashwin Kalyan¹ Peter Anderson¹ Stefan Lee¹ Dhruv Batra^{1,2}

Split	Facility Location Accuracy	
	CIDEr	SPICE
(submodular)	1.5995	0.1745
(arbitrary)	1.5324	0.1723

Table 1. Constraining the subset selector to be submodular promotes a favorable inductive bias as observed by the deterioration in performance when using arbitrary function to obtain utility.

Using Non-submodular Functions. As discussed in Section 2.3, not restricting the subset selection function f to be submodular, reduces our approach to a *trainable* variant of DivMBest. While this strategy allows for more complex functions to be learnt, it lacks the theoretical guarantees of submodular maximization and can potentially make the learning hard; owing to a larger model class.

¹School of Interactive Computing, Georgia Tech, Atlanta, GA, USA ²Facebook AI Research, Menlo Park, CA, USA. Correspondence to: Ashwin Kalyan <ashwinkv@gatech.edu>.