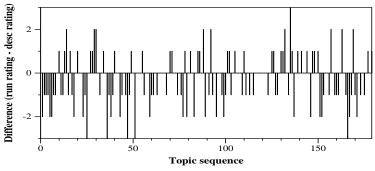| Summary Statistics	
Run ID	udel_wang_zheng3
Task	Summarization
Run type	automatic
Summarization model	Supervised neural or deep learning model; Transfer learning or other pre-trained model
Summarization method	Abstractive
Number of summaries	179

Overall measures	
Summaries rated "excellent"	10 (5.6%)
Summaries rated "good"	34 (19.0%)
Summaries rated "fair"	86 (48.0%)
Summaries rated "bad"	49 (27.4%)

ROUGE score compared to episode description	
ROUGE_L P score	0.2084 (95%-conf.int. 0.19976 - 0.21738)
ROUGE_L R score	0.1601 (95%-conf.int. 0.15262 - 0.16751)
ROUGE_L F score	0.1529 (95%-conf.int. 0.14714 - 0.15894)

Qualitative judgments	
Q1: Does the summary include names of the main people (hosts, guests, characters) involved or mentioned in the podcast?	52.0%
Q2: Does the summary give any additional information about the people mentioned (such as their job titles, biographies, personal background, etc)?	30.7%
Q3: Does the summary include the main topic(s) of the podcast?	66.5%
Q4: Does the summary tell you anything about the format of the podcast; e.g. whether it's an interview, whether it's a chat between friends, a monologue, etc?	52.0%
Q5: Does the summary give you mode context on the title of the podcast?	58.1%
Q6: Does the summary contain redundant information?	8.9%
Q7: Is the summary written in good English?	62.6%
Q8: Are the start and end of the summary good sentence and paragraph start and end points?	46.4%

Per-topic difference in rating from episode description