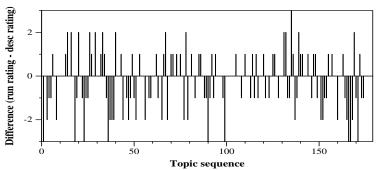| Summary Statistics	
Run ID	udel_wang_zheng4
Task	Summarization
Run type	automatic
Summarization model	Supervised neural or deep learning model; Transfer learning or other pre-trained model
Summarization method	Abstractive
Number of summaries	179

Overall measures	
Summaries rated "excellent"	14 (7.8%)
Summaries rated "good"	39 (21.8%)
Summaries rated "fair"	74 (41.3%)
Summaries rated "bad"	52 (29.1%)

ROUGE score compared to episode description	
ROUGE_L P score	0.2016 (95%-conf.int. 0.19287 - 0.21045)
ROUGE_L R score	0.1683 (95%-conf.int. 0.16077 - 0.17618)
ROUGE_L F score	0.1549 (95%-conf.int. 0.14899 - 0.16043)

Qualitative judgments	
Q1: Does the summary include names of the main people (hosts, guests, characters) involved or mentioned in the podcast?	56.4%
Q2: Does the summary give any additional information about the people mentioned (such as their job titles, biographies, personal background, etc)?	28.5%
Q3: Does the summary include the main topic(s) of the podcast?	68.2%
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?	63.1%
Q5: Does the summary give you mode context on the title of the podcast?	62.0%
Q6: Does the summary contain redundant information?	4.5%
Q7: Is the summary written in good English?	74.9%
Q8: Are the start and end of the summary good sentence and paragraph start and end points?	51.4%

Per-topic difference in rating from episode description