

Podcasts Track results — (podcast_baselines) Baseline runs for Podcast track

Summary Statistics	
Run ID	bartcnn
Task	Summarization
Run type	automatic
Summarization 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"	25 (14.0%)
Summaries rated "fair"	88 (49.2%)
Summaries rated "bad"	56 (31.3%)

ROUGE score compared to episode description	
ROUGE_L P score	0.0849 (95%-conf.int. 0.08071 - 0.08973)
ROUGE_L R score	0.2719 (95%-conf.int. 0.26190 - 0.28163)
ROUGE_L F score	0.1130 (95%-conf.int. 0.10834 - 0.11784)

Qualitative judgments	
Q1: Does the summary include names of the main people (hosts, guests, characters) involved or mentioned in the podcast?	72.6%
Q2: Does the summary give any additional information about the people mentioned (such as their job titles, biographies, personal background, etc)?	48.0%
Q3: Does the summary include the main topic(s) of the podcast?	69.3%
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?	58.1%
Q5: Does the summary give you more context on the title of the podcast?	67.6%
Q6: Does the summary contain redundant information?	25.7%
Q7: Is the summary written in good English?	34.6%
Q8: Are the start and end of the summary good sentence and paragraph start and end points?	22.3%

