Podcasts Track results — (spotify) Spotify

Summary Statistics	
Run ID	coarse2fine
Task	Summarization
Run type	automatic
Summarization model	Graph model; Transfer learning or other pre-trained model
Summarization method	Abstractive
Number of summaries	179

Overall measures	
Summaries rated "excellent"	18 (10.1%)
Summaries rated "good"	39 (21.8%)
Summaries rated "fair"	53 (29.6%)
Summaries rated "bad"	69 (38.5%)

ROUGE score compared to episode description	
ROUGE_L P score	0.1576 (95%-conf.int. 0.14914 - 0.16610)
ROUGE_L R score	0.1875 (95%-conf.int. 0.17972 - 0.19545)
ROUGE_L F score	0.1359 (95%-conf.int. 0.13080 - 0.14145)

Qualitative judgments	
Q1: Does the summary include names of the main people (hosts, guests, characters) involved or mentioned in the podcast?	35.2%
Q2: Does the summary give any additional information about the people mentioned (such as their job titles, biographies, personal background, etc)?	22.9%
Q3: Does the summary include the main topic(s) of the podcast?	60.3%
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?	41.3%
Q5: Does the summary give you mode context on the title of the podcast?	57.0%
Q6: Does the summary contain redundant information?	10.1%
Q7: Is the summary written in good English?	79.9%
Q8: Are the start and end of the summary good sentence and paragraph start and end points?	55.3%

Per-topic difference in rating from episode description