Podcasts Track results — (UTDThesis) Mathew Perez UT Dallas Graudate Thesis

Summary Statistics	
Run ID	UTDThesis1
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"	1 (0.6%)
Summaries rated "good"	10 (5.6%)
Summaries rated "fair"	53 (29.6%)
Summaries rated "bad"	115 (64.2%)

ROUGE score compared to episode description	
ROUGE_L P score	0.1716 (95%-conf.int. 0.16477 - 0.17810)
ROUGE_L R score	0.1292 (95%-conf.int. 0.12344 - 0.13502)
ROUGE_L F score	0.1249 (95%-conf.int. 0.12067 - 0.12922)

Qualitative judgments	
Q1: Does the summary include names of the main people (hosts, guests, characters) involved or mentioned in the podcast?	30.2%
Q2: Does the summary give any additional information about the people mentioned (such as their job titles, biographies, personal background, etc)?	17.3%
Q3: Does the summary include the main topic(s) of the podcast?	36.9%
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?	48.0%
Q5: Does the summary give you mode context on the title of the podcast?	31.3%
Q6: Does the summary contain redundant information?	7.3%
Q7: Is the summary written in good English?	49.2%
Q8: Are the start and end of the summary good sentence and paragraph start and end points?	14.0%

Per-topic difference in rating from episode description