Podcasts Track results — (UCF_NLP) University of Central Florida

Summary Statistics	
Run ID	UCF_NLP1
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"	34 (19.0%)
Summaries rated "good"	45 (25.1%)
Summaries rated "fair"	68 (38.0%)
Summaries rated "bad"	32 (17.9%)

ROUGE score compared to episode description	
ROUGE_L P score	0.2667 (95%-conf.int. 0.25302 - 0.28002)
ROUGE_L R score	0.2205 (95%-conf.int. 0.21003 - 0.23113)
ROUGE_L F score	0.2132 (95%-conf.int. 0.20307 - 0.22333)

Qualitative judgments	
Q1: Does the summary include names of the main people (hosts, guests, characters) involved or mentioned in the podcast?	60.9%
Q2: Does the summary give any additional information about the people mentioned (such as their job titles, biographies, personal background, etc)?	37.4%
Q3: Does the summary include the main topic(s) of the podcast?	80.4%
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?	56.4%
Q5: Does the summary give you mode context on the title of the podcast?	72.1%
Q6: Does the summary contain redundant information?	7.8%
Q7: Is the summary written in good English?	82.7%
Q8: Are the start and end of the summary good sentence and paragraph start and end points?	65.9%

Per-topic difference in rating from episode description