

Podcasts Track results — (hk_uu_podcast) Uppsala University

Summary Statistics	
Run ID	hk_uu_podcast1
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"	35 (19.6%)
Summaries rated "good"	54 (30.2%)
Summaries rated "fair"	64 (35.8%)
Summaries rated "bad"	26 (14.5%)

ROUGE score compared to episode description	
ROUGE_L P score	0.2646 (95%-conf.int. 0.25362 - 0.27536)
ROUGE_L R score	0.1898 (95%-conf.int. 0.18083 - 0.19886)
ROUGE_L F score	0.1923 (95%-conf.int. 0.18421 - 0.20084)

Qualitative judgments	
Q1: Does the summary include names of the main people (hosts, guests, characters) involved or mentioned in the podcast?	58.1%
Q2: Does the summary give any additional information about the people mentioned (such as their job titles, biographies, personal background, etc)?	34.6%
Q3: Does the summary include the main topic(s) of the podcast?	82.1%
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?	57.0%
Q5: Does the summary give you more context on the title of the podcast?	73.2%
Q6: Does the summary contain redundant information?	6.7%
Q7: Is the summary written in good English?	91.1%
Q8: Are the start and end of the summary good sentence and paragraph start and end points?	72.1%

