SummScreen: A Dataset for Abstractive Screenplay Summarization

Mingda Chen, Zewei Chu, Sam Wiseman, Kevin Gimpel

Duke
UNIVERSITY

SummScreen

- An abstractive summarization dataset combining TV series transcripts and episode recaps.
- SummScreen is constructed from fan-contributed websites.

Transcript:

[The apartment]
Sheldon : What color would you like to be ?
Leonard : Well , I 'd like to be green , but you know you always take it .
Sheldon : That 's not true . Any color 's fine with me . Yeah , I could be a - a combination of blue and yellow .
Leonard : Blue and yellow make green .
Sheldon : Well , then it 's settled .

Penny : Hi . Ready to go ?
Sheldon : Oh , good news , we ordered lunch , so we can all stay here and play Lord of the Rings Risk .
Amy : Sheldon , we said that we would play games with you tonight .
Sheldon : Oh , no , we 'll still be playing it tonight , this game can easily take eight hours .

Penny : Sweetie , you really thought I 'd want to do this ?
Leonard : No .
Penny : Well , did you tell him that ?
Leonard : Yes .
Penny : Did you say it out loud with words ?
Leonard : No .
Penny : I do n't want to spend the whole day playing a board game .
…

Recap:

Sheldon and Leonard are happy playing a board game until Amy and Penny say they are tired of doing what the guys want …

SummScreen

- SummScreen has
  - ≈ 100 TV shows
  - ≈ 30k episodes
  - ≈ 300~400 lines for each transcript
  - ≈ 300 word tokens for each recap

Transcript:

[The apartment]
Sheldon : What color would you like to be ?
Leonard : Well , I 'd like to be green , but you know you always take it .
Sheldon : That 's not true . Any color 's fine with me . Yeah , I could be a - a combination of blue and yellow .
Leonard : Blue and yellow make green .
Sheldon : Well , then it 's settled .

Penny : Hi . Ready to go ?
Sheldon : Oh , good news , we ordered lunch , so we can all stay here and play Lord of the Rings Risk .
Amy : Sheldon , we said that we would play games with you tonight .
Sheldon : Oh , no , we 'll still be playing it tonight , this game can easily take eight hours .

Penny : Sweetie , you really thought I 'd want to do this ?
Leonard : No .
Penny : Well , did you tell him that ?
Leonard : Yes .
Penny : Did you say it out loud with words ?
Leonard : No .
Penny : I do n't want to spend the whole day playing a board game .
…

Recap:

Sheldon and Leonard are happy playing a board game until Amy and Penny say they are tired of doing what the guys want …

Dataset Comparison

	#instances	#tokens(input)	#tokens(summary)	#speakers	Domain
SAMSum	16.4k	83.9	20.3	2.2	Chitchat
Forumsum	4.1k	303.5	36.0	6.7	Forum Messages
MediaSum	463.6k	1.6k	14.4	6.5	News Interviews
QMSum	1.8k	9.1k	69.6	9.2	Meetings
SummScreen	26.9k	6.6k	337.4	28.3	TV Series

Dataset Comparison

	#instances	#tokens(input)	#tokens(summary)	#speakers	Domain
SAMSum	16.4k	83.9	20.3	2.2	Chitchat
Forumsum	4.1k	303.5	36.0	6.7	Forum Messages
MediaSum	463.6k	1.6k	14.4	6.5	News Interviews
QMSum	1.8k	9.1k	69.6	9.2	Meetings
SummScreen	26.9k	6.6k	337.4	28.3	TV Series

SummScreen combines long source inputs, large numbers of speakers, and a moderate number of instances.

Dataset Comparison

	#instances	#tokens(input)	#tokens(summary)	#speakers	Domain
SAMSum	16.4k	83.9	20.3	2.2	Chitchat
Forumsum	4.1k	303.5	36.0	6.7	Forum Messages
MediaSum	463.6k	1.6k	14.4	6.5	News Interviews
QMSum	1.8k	9.1k	69.6	9.2	Meetings
SummScreen	26.9k	6.6k	337.4	28.3	TV Series

SummScreen combines long source inputs, large numbers of speakers, and a moderate number of instances.

Dataset Comparison

	#instances	#tokens(input)	#tokens(summary)	#speakers	Domain
SAMSum	16.4k	83.9	20.3	2.2	Chitchat
Forumsum	4.1k	303.5	36.0	6.7	Forum Messages
MediaSum	463.6k	1.6k	14.4	6.5	News Interviews
QMSum	1.8k	9.1k	69.6	9.2	Meetings
SummScreen	26.9k	6.6k	337.4	28.3	TV Series

SummScreen combines long source inputs, large numbers of speakers, and a moderate number of instances.

Dataset Comparison

	#instances	#tokens(input)	#tokens(summary)	#speakers	Domain
SAMSum	16.4k	83.9	20.3	2.2	Chitchat
Forumsum	4.1k	303.5	36.0	6.7	Forum Messages
MediaSum	463.6k	1.6k	14.4	6.5	News Interviews
QMSum	1.8k	9.1k	69.6	9.2	Meetings
SummScreen	26.9k	6.6k	337.4	28.3	TV Series

the unique domain leads to characteristics: e.g.,

- complex interactions between characters
- character dialogues + action descriptions

Dataset Challenges

- Plot details are often expressed indirectly in character dialogues
  - References about a board game (shown in red)

Transcript:

[The apartment]
Sheldon : What color would you like to be ?
Leonard : Well , I 'd like to be green , but you know you always take it .
Sheldon : That 's not true . Any color 's fine with me . Yeah , I could be a - a combination of blue and yellow .
Leonard : Blue and yellow make green .
Sheldon : Well , then it 's settled .

Penny : Hi . Ready to go ?
Sheldon : Oh , good news , we ordered lunch , so we can all stay here and play Lord of the Rings Risk .
Amy : Sheldon , we said that we would play games with you tonight .
Sheldon : Oh , no , we 'll still be playing it tonight , this game can easily take eight hours .

Penny : Sweetie , you really thought I 'd want to do this ?
Leonard : No .
Penny : Well , did you tell him that ?
Leonard : Yes .
Penny : Did you say it out loud with words ?
Leonard : No .
Penny : I do n't want to spend the whole day playing a board game .
…

Recap:

Sheldon and Leonard are happy playing a board game until Amy and Penny say they are tired of doing what the guys want …

Dataset Challenges

- Plot details are often expressed indirectly in character dialogues
  - References about a board game (shown in red)
  - Clues about the characters' feelings about playing the board game (underlined)

Transcript:

[The apartment]
Sheldon : What color would you like to be ?
Leonard : Well , I 'd like to be green , but you know you always take it .
Sheldon : That 's not true . Any color 's fine with me . Yeah , I could be a - a combination of blue and yellow .
Leonard : Blue and yellow make green .
Sheldon : Well , then it 's settled .

Penny : Hi . Ready to go ?
Sheldon : Oh , good news , we ordered lunch , so we can all stay here and play Lord of the Rings Risk .
Amy : Sheldon , we said that we would play games with you tonight .
Sheldon : Oh , no , we 'll still be playing it tonight , this game can easily take eight hours .

Penny : Sweetie , you really thought I 'd want to do this ?
Leonard : No .
Penny : Well , did you tell him that ?
Leonard : Yes .
Penny : Did you say it out loud with words ?
Leonard : No .
Penny : I do n't want to spend the whole day playing a board game .
…

Recap:

Sheldon and Leonard are happy playing a board game until Amy and Penny say they are tired of doing what the guys want …

Dataset Challenges

- Plot details are often expressed indirectly in character dialogues
- Plot details may be scattered across the entirety of the transcript

Transcript:

Line 119 DOCTOR : Camera ! Camera ! (takes camera from ALEC 'S unresisting hands)

…

Line 212 The DOCTOR turns around and continues to take photos with the camera …

…

Line 287 The DOCTOR steps out of the TARDIS wearing the spacesuit … He scans with the sonic before picking up the camera to take a few pictures .

…

Line 336 DOCTOR : Right ! Done ! That 's it … She 's not a ghost ... but she 's definitely a lost soul . (walks over to screen) Her name 's Hila Tacorian . She 's a pioneer , a time traveller - or at least she will be , in a few hundred years .

Recap:

… the Doctor borrows Alec 's camera and uses the TARDIS to take pictures of the mansion 's location throughout time . Thanks to this , the Doctor learns it 's not a ghost in the pictures , but a time traveler named Hila Tacorian …

Dataset Challenges

- Plot deta
  indirectl
- Plot details may be scattered across the entirety of the transcript

Line numbers in the original transcript

Transcript:

Line 119 DOCTOR : Camera ! Camera ! (takes camera from ALEC 'S unresisting hands)

…

Line 212 The DOCTOR turns around and continues to take photos with the camera …

…

Line 287 The DOCTOR steps out of the TARDIS wearing the spacesuit … He scans with the sonic before picking up the camera to take a few pictures .

…

Line 336 DOCTOR : Right ! Done ! That 's it … She 's not a ghost … but she 's definitely a lost soul . (walks over to screen) Her name 's Hila Tacorian . She 's a pioneer , a time traveller - or at least she will be , in a few hundred years .

Recap:

… the Doctor borrows Alec 's camera and uses the TARDIS to take pictures of the mansion 's location throughout time . Thanks to this , the Doctor learns it 's not a ghost in the pictures , but a time traveler named Hila Tacorian …

Methods

- Neural models
- Nearest neighbor models (NNM): retrieve summaries from the training set
- Hybrid models (nearest neighbor content selector → neural models)

Evaluation Metrics

- Generic metrics: BLEU and ROUGE scores
- Entity metrics:
  - Bag of characters (BoC): the fraction of the **characters** overlapping with gold
  - Bag of character relations (BoR): the fraction of the **cooccurred character pairs** overlapping with gold

Experimental Results

	Generic metrics	Entity metrics
NNM	**18.6**	43.3
Neural models	14.1	32.8
Hybrid	13.7	41.5
Hybrid (w/ oracle content selector)	14.4	**48.5**

Average scores for the generic metrics and entity metrics when evaluating the models on part of the SummScreen test set

Experimental Results

	Generic metrics	Entity metrics
NNM	**18.6**	43.3
Neural models	14.1	32.8
Hybrid	13.7	41.5
Hybrid (w/ oracle content selector)	14.4	**48.5**

Average scores for the generic metrics and entity metrics when evaluating the models on part of the SummScreen test set

- NNM shows strong performance.

Experimental Results

	Generic metrics	Entity metrics
NNM	**18.6**	43.3
Neural models	14.1	32.8
Hybrid	13.7	41.5
Hybrid (w/ oracle content selector)	14.4	**48.5**

Average scores for the generic metrics and entity metrics when evaluating the models on part of the SummScreen test set

- NNM shows strong performance.
- With the help of the oracle content selector, the hybrid model improves significantly in both semantic matching and entity-related metrics.

Conclusion

- We constructed an abstractive summarization dataset SummScreen from fan-contributed websites.
- SummScreen has a unique set of challenges: drawing information from a wide range of the input and understanding the context, among others.
- Having a strong content selector can benefit model performance on SummScreen.

Conclusion

- SummScreen is used in
  - a benchmark for long text (SCROLLS)
  - shared tasks for an upcoming workshop (Creative-Summ)
- Scan the QR code to check out our dataset!