TATA CONSULTANCY SERVICES
Experience certainty.

TATA

Topic oriented Summarization using Transformer based Encoder Decoder model

Vipul Chauhan, Prasenjeet Roy, Lipika Dey and Tushar Goel

TCS Research
New Delhi, India
(chauhan.vipul, r.prasenjeet2, Lipika.dey, t.goel)@tcs.com

DialogSum Challenge: Summarizing Real-Life Scenario Dialogues[1]

- Challenge – Given a piece of multi-turn dialogue text, generate a salient, concise, fluent and coherent summary
- Motivation - investigate the challenges in dialogue summarization.
- Dialogue – an important channel for achieving communication intents, different from monologic texts in nature
- Applications - chatbot and personal assistant.

Dialogue Text:
#Person1#: Why didn't you tell me you had a girlfriend?
#Person2#: Sorry, I thought you knew.
#Person1#: But you should tell me you were in love with her.
#Person2#: Didn't I?
#Person1#: You know you didn't.
#Person2#: Well, I am telling you now.
#Person1#: Yes, but you might have told me before.
#Person2#: I didn't think you would be interested.
#Person1#: You can't be serious. How dare you not tell me you are going to marry her?
#Person2#: Sorry, I didn't think it mattered.
#Person1#: Oh, you men! You are all the same.

Summary from Dataset:
#Person1#'s angry because **#Person2#** didn't tell **#Person1#** that **#Person2#** had a girlfriend and would marry her.
Topic from Dataset: have a girlfriend

Dataset Description

Public dataset - train, dev, test and hidden-test dataset[2]

Train - (12,460), dev - (500) consist of single human annotated summary and short topic for each dialogue

Test - (500) consists of 3 annotated summary and associated 3 short topic for each dialogue.

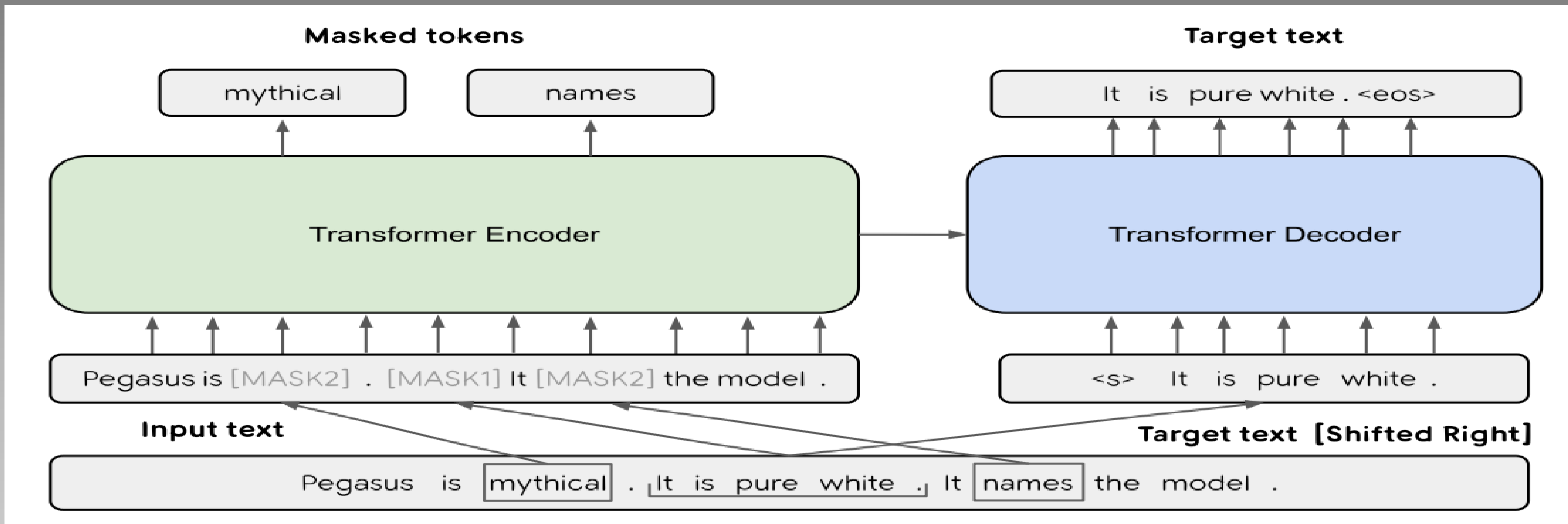
Hidden test - (100) dialogue and topics.

Challenges

- Diverse Interactive patterns between speakers as well as inherent topic drifts
- A frequent coreference and ellipsis make dialogue difficult to understand.
- Linguistic phenomena make dialogues difficult to encode using ordinary representation learning technologies.
- Objective summary requires summary from an observer's perspective
- Dialogue actions at the pragmatic level – It not only summarizes what the speakers are saying but also what they are doing

PEGASUS[3]

Pre-training with Extracted Gap-sentences for Abstractive Summarization Sequence-to-sequence models



The base architecture of PEGASUS [3]. Both GSG and MLM are applied simultaneously in this example.

Variants –

- $\text{PEGASUS}_{\text{BASE}}$ – 12 layers and H=768
- $\text{PEGASUS}_{\text{LARGE}}$ – 16 layers and H=1024

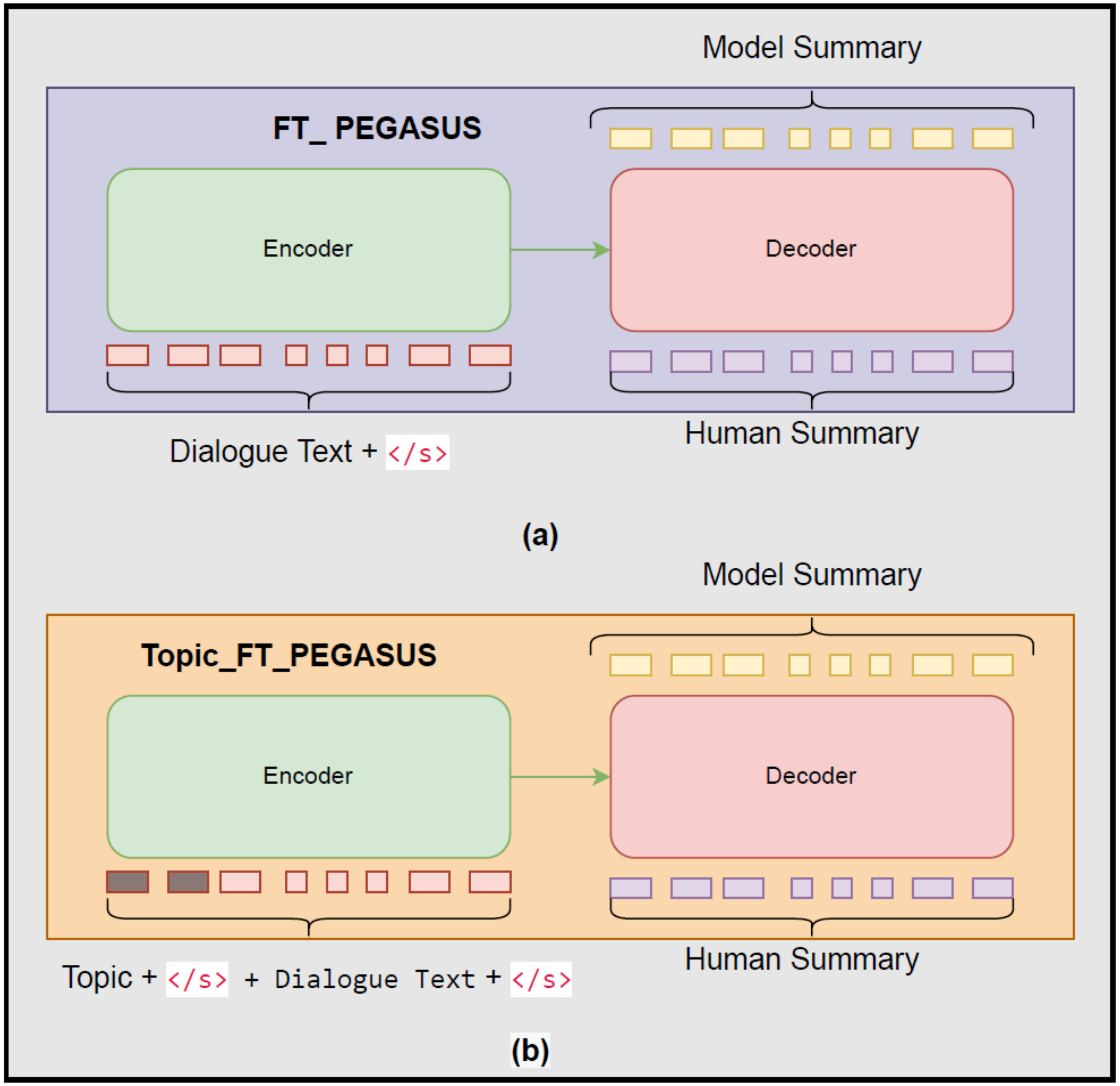
Pre-training

Corpora – C4 & HugeNews

Objectives –

- Masked Language Model (MLM)
- Gap Sentence Generation (GSG)

Proposed Method – Topic_FT_PEGASUS



Sample Output

Dialogue Text

#Person1#: Who stands out in your mind as a man or woman of sound character?
#Person2#: If I think of famous people, I think of Abraham Lincoln.
#Person1#: He's the US president, who walked five miles just to give a lady her change, isn't he?
#Person2#: That's the one. He also was famous for never giving up on his goals.
#Person1#: That's right. He ran for office quite a few times before he was finally elected.
#Person2#: And I also admire him for his courage in fighting for equal rights.
#Person1#: He had great vision, didn't he?
#Person2#: And humility. I would have liked to meet him personally.

Topic – sound character

Model Summary	Human Summary
#Person1# and #Person2# talk about who stands out in their mind as a man or woman of sound character.	#Person1# and #Person2# are talking about Abraham Lincoln. They think he was a noble man.

Topic – famous people

Model Summary	Human Summary
#Person1# and #Person2# are talking about famous people. They admire Abraham Lincoln for his great vision, courage, and humility.	#Person2# admires Abraham Lincoln for his perseverance, courage and humility.

Topic – discuss Abraham Lincoln

Model Summary	Human Summary
#Person1# and #Person2# talk about Abraham Lincoln as a man or woman of sound character.	#Person1# and #Person2# talk about Abraham Lincoln and his glorious history. They both admire him.

Experiments and Results

Model	Average Score				Best Score			
	R1	R2	RL	B-S	R1	R2	RL	B-S
PT_PEGASUS	25.99	6.41	20.97	87.77	37.63	9.63	26.48	88.15
FT_PEGASUS	43.36	18.36	36.23	92.19	51.59	26.58	45.54	92.64
Topic_FT_PEGASUS	**49.42**	**21.81**	**40.85**	**92.22**	**54.53**	**32.00**	**51.47**	**93.22**

Evaluation results over the public test dataset.

Baselines

- PT_PEGASUS – pretrained PEGASUS_LARGE generate the summary using dialogue text as an input.
- FT_PEGASUS – fine-tuned PEGASUS_LARGE using DialogSum train and dev dataset with input dialogue text as input.

Evaluation Metric - ROUGE and BERTScore

Observations

- Fine-tuning helped the model to learn the linguistic phenomena of dialogues.
- Incorporating topics while fine-tuning allows the model to focus on different text segments centred around a given topic.

DialogSum Challenge Leaderboard

S.No.	Team	R1	R2	RL	BERTScore
1	GoodBai	49.66	**26.03**	48.44	91.69
2	UoT	49.75	25.15	46.50	91.76
3	IITP-CUNI	45.89	21.88	43.16	91.13
4	TCS_WITM_2022*	**50.32**	25.59	47.40	**91.81**

Evaluation Result on the Hidden Testset. Evaluated by the challenge organisers.

*evaluated the prediction topic-wise

Conclusions and Future Work

- Generate a topic-oriented summary using pre-trained abstractive model
- Shown that the PEGASUS can be fine-tuned using the proposed methodology.
- Proposed methodology performed significantly better than baselines.

Future work

- Incorporating nuances of dialogue, speech act theory etc.
- Propose a model that generate topic from dialogue

References

[1] Yulong Chen, Yang Liu, and Yue Zhang. 2021b. Dialogsum challenge: Summarizing real-life scenario dialogues. In Proceedings of the 14th International Conference on Natural Language Generation, pages 308–313.

[2] Yulong Chen, Yang Liu, Liang Chen, and Yue Zhang. 2021a. DialogSum: A real-life scenario dialogue summarization dataset. In Findings of the Association for Computational Linguistics: ACL-IJCNLP 2021, pages 5062–5074, Online. Association for Computational Linguistics.

[3] Jingqing Zhang, Yao Zhao, Mohammad Saleh, and Peter Liu. 2020. Pegasus: Pre-training with extracted gap-sentences for abstractive summarization. In International Conference on Machine Learning, pages 11328–11339. PMLR.

TATA CONSULTANCY SERVICES
Experience certainty.

Thank You

Vipul Chauhan
TCS Research, Tata Consultancy Services
New Delhi, India

chauhan.vipul@tcs.com