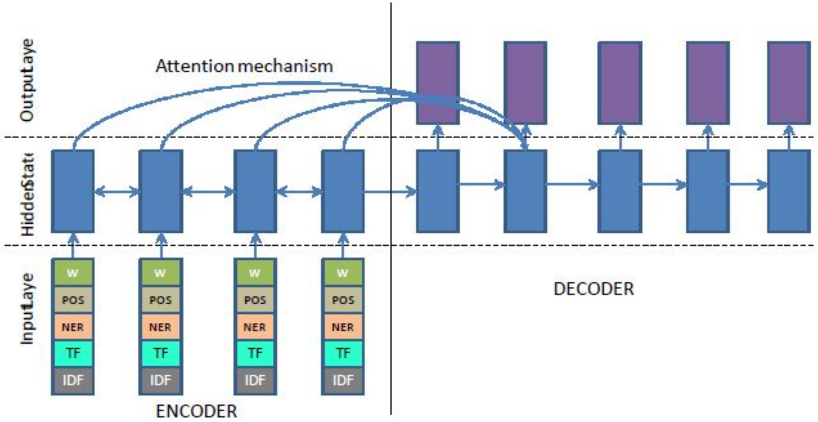
Pretraining-Based Natural Language Generation for Text Summarization

Yixian Liu
2019/4/17

Text summarization

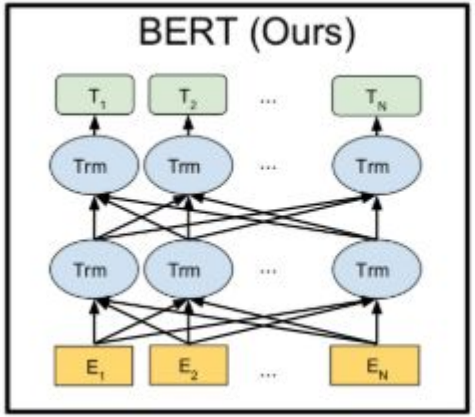
seq-to-seq model

from a document to its summarization

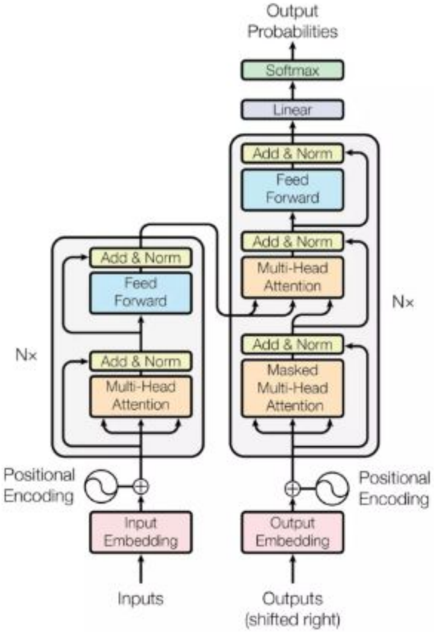


Bert and transformer

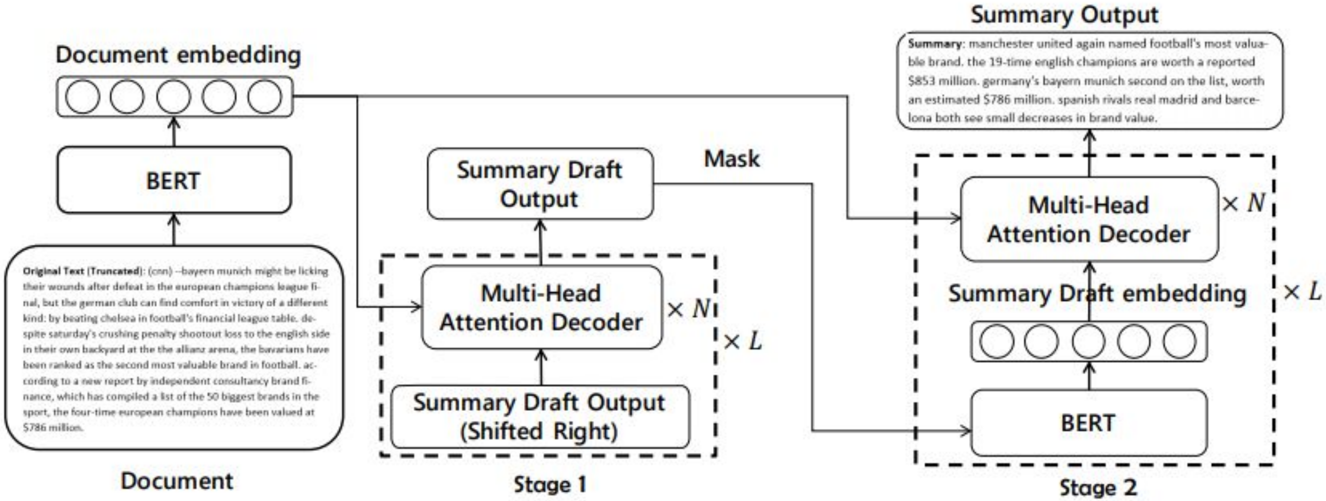
Bert



Transformer



Pretraining-Based model



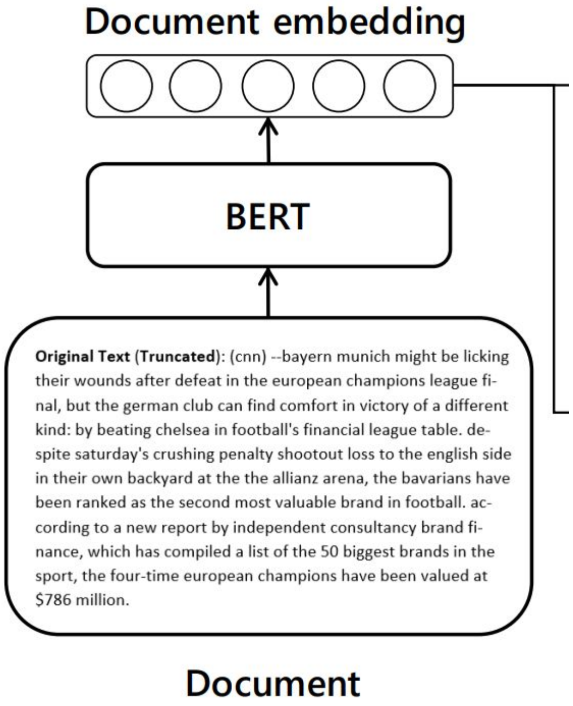
Summary: manchester united again named football's most valuable brand. the 19-time english champions are worth a reported $853 million. germany's bayern munich second on the list, worth an estimated $786 million. spanish rivals real madrid and barcelona both see small decreases in brand value.

Original Text (Truncated): (cnn) --bayern munich might be licking their wounds after defeat in the european champions league final, but the german club can find comfort in victory of a different kind: by beating chelsea in football's financial league table. despite saturday's crushing penalty shootout loss to the english side in their own backyard at the the allianz arena, the bavarians have been ranked as the second most valuable brand in football. according to a new report by independent consultancy brand finance, which has compiled a list of the 50 biggest brands in the sport, the four-time european champions have been valued at $786 million.

Encoder

$$H = BERT(x_1, \ldots, x_m)$$

$$H = \{h_1, \ldots, h_m\}$$



Original Text (Truncated): (cnn) --bayern munich might be licking their wounds after defeat in the european champions league final, but the german club can find comfort in victory of a different kind: by beating chelsea in football's financial league table. despite saturday's crushing penalty shootout loss to the english side in their own backyard at the the allianz arena, the bavarians have been ranked as the second most valuable brand in football. according to a new report by independent consultancy brand finance, which has compiled a list of the 50 biggest brands in the sport, the four-time european champions have been valued at $786 million.

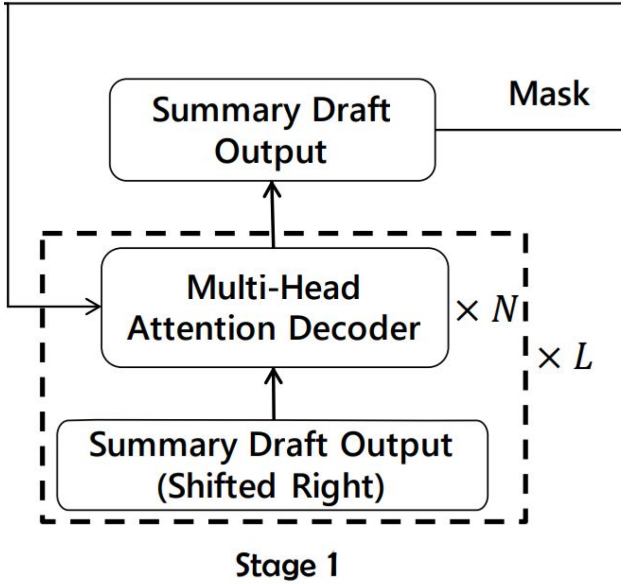
Stage I decoder

Time step t

$$P_t^{vocab}(w) = f_{dec}(q_{<t}, H)$$
$$L_{dec} = \sum_{t=1}^{|a|} - \log P(a_t = y_t^* | a_{<t}, H)$$

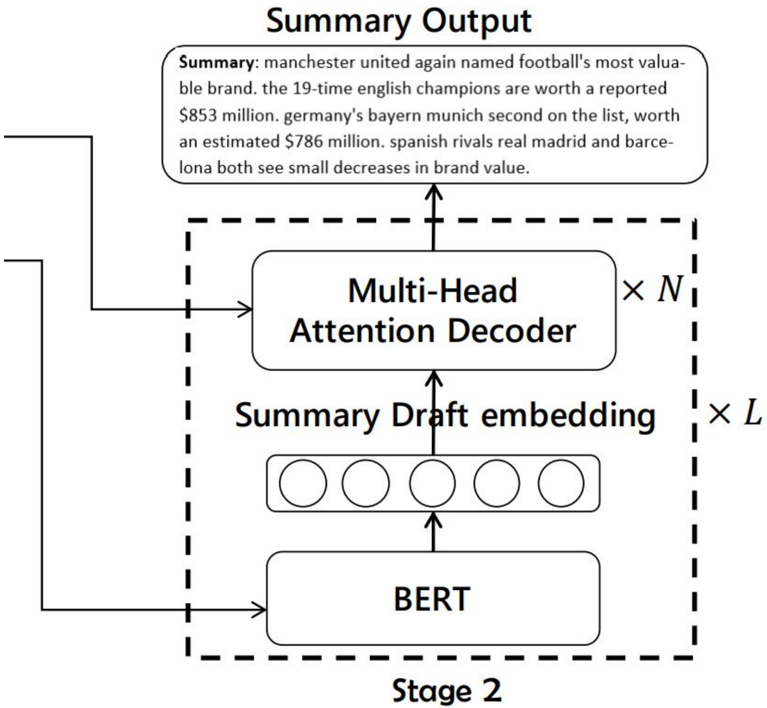
Copy mechanism

$$P_t(w) = (1 - g_t) P_t^{vocab}(w) + g_t \sum_{i:w_i = w} \alpha_t^i$$


Mask
Summary Draft Output
Multi-Head Attention Decoder
$\times N$
$\times L$
Summary Draft Output (Shifted Right)
Stage 1

Stage II decoder ---- Summary Refine

$$L_{refine} = \sum_{t=1}^{|y|} - \log P(y_t = y_t^* | a_{\neq t}, H)$$



Summary: manchester united again named football's most valuable brand. the 19-time english champions are worth a reported $853 million. germany's bayern munich second on the list, worth an estimated $786 million. spanish rivals real madrid and barcelona both see small decreases in brand value.

Mixed objective

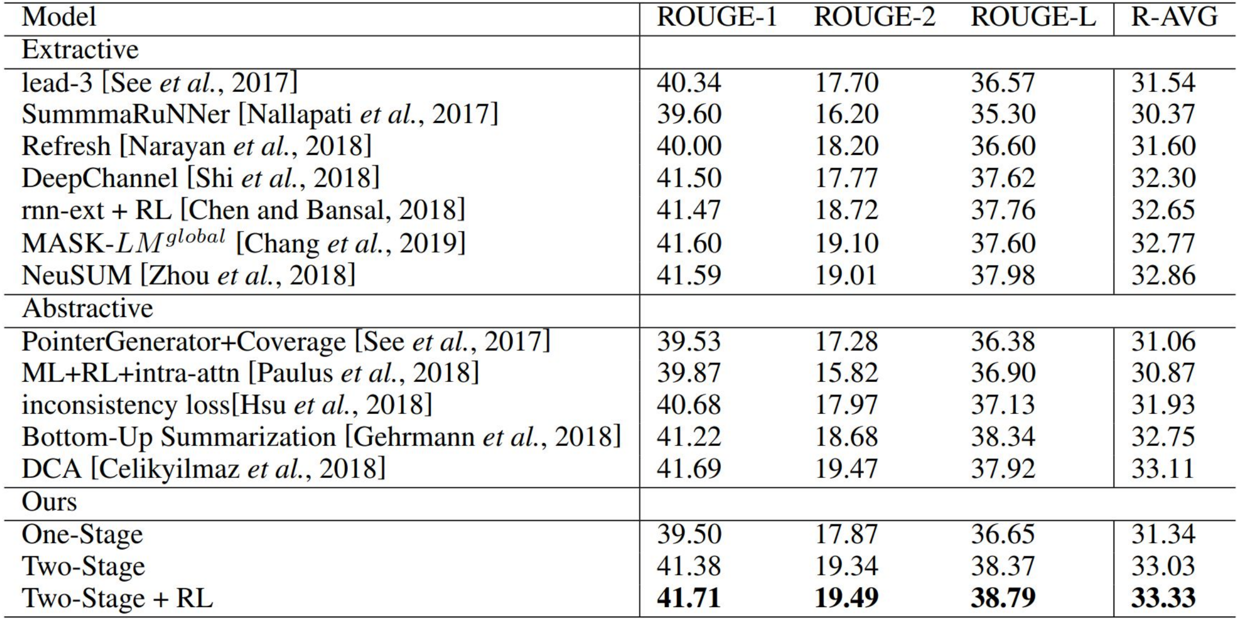
Policy gradient about ROUGE

$$L_{dec}^{rl} = R(a^s) \cdot [-\log(P(a^s|x))]$$
$$\hat{L}_{dec} = \gamma * L_{dec}^{rl} + (1-\gamma) * L_{dec}$$

Final objective function

$$L_{model} = \hat{L}_{dec} + \hat{L}_{refine}$$

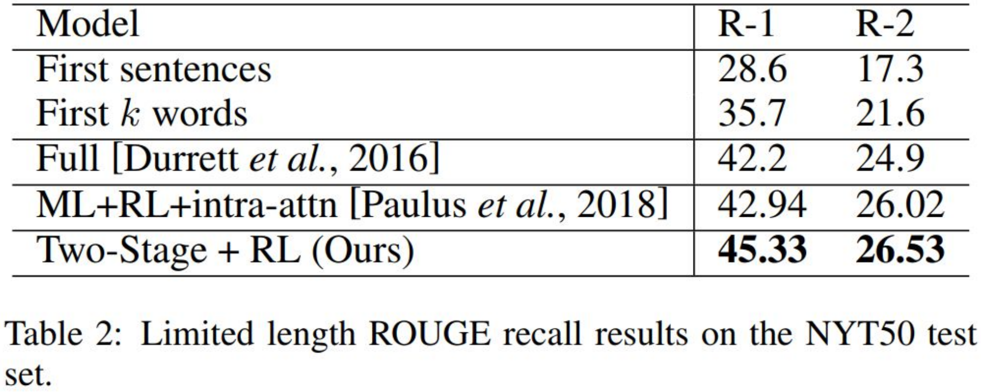
Result



Model	ROUGE-1	ROUGE-2	ROUGE-L	R-AVG
Extractive				
lead-3 [See *et al.*, 2017]	40.34	17.70	36.57	31.54
SummmaRuNNer [Nallapati *et al.*, 2017]	39.60	16.20	35.30	30.37
Refresh [Narayan *et al.*, 2018]	40.00	18.20	36.60	31.60
DeepChannel [Shi *et al.*, 2018]	41.50	17.77	37.62	32.30
rnn-ext + RL [Chen and Bansal, 2018]	41.47	18.72	37.76	32.65
MASK-LM^{global} [Chang *et al.*, 2019]	41.60	19.10	37.60	32.77
NeuSUM [Zhou *et al.*, 2018]	41.59	19.01	37.98	32.86
Abstractive				
PointerGenerator+Coverage [See *et al.*, 2017]	39.53	17.28	36.38	31.06
ML+RL+intra-attn [Paulus *et al.*, 2018]	39.87	15.82	36.90	30.87
inconsistency loss[Hsu *et al.*, 2018]	40.68	17.97	37.13	31.93
Bottom-Up Summarization [Gehrmann *et al.*, 2018]	41.22	18.68	38.34	32.75
DCA [Celikyilmaz *et al.*, 2018]	41.69	19.47	37.92	33.11
Ours				
One-Stage	39.50	17.87	36.65	31.34
Two-Stage	41.38	19.34	38.37	33.03
Two-Stage + RL	**41.71**	**19.49**	**38.79**	**33.33**

Table 1: ROUGE F1 results for various models and ablations on the CNN/Daily Mail test set. R-AVG calculates average score of Rouge-1, Rouge-2 and Rouge-L.

Result



Model	R-1	R-2
First sentences	28.6	17.3
First k words	35.7	21.6
Full [Durrett *et al.*, 2016]	42.2	24.9
ML+RL+intra-attn [Paulus *et al.*, 2018]	42.94	26.02
Two-Stage + RL (Ours)	**45.33**	**26.53**

Table 2: Limited length ROUGE recall results on the NYT50 test set.