A Case Study on Neural Headline Generation for Editing Support

Kazuma Murao[*,1], Ken Kobayashi[*,1], **Hayato Kobayashi[1,2]**, Taichi Yatsuka[1], Takeshi Masuyama[1], Tatsuru Higurashi[1], Yoshimune Tabuchi[1]

[1]Yahoo Japan Corporation [2]RIKEN AIP

([*]Equal contribution)

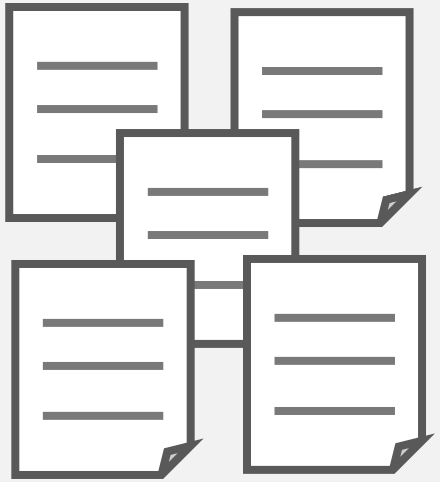
Summary

- Our work
  - Address “short title” generation for a news aggregation service, where editors create short titles to introduce important articles
- Contributions
  - Show a practical use case of neural headline generation
    - Most news articles basically already have headlines
  - Propose an encoder-decoder model with multiple encoders
  - Deploy our model to an editing support tool and show the results of comparing the editors’ behavior

Yahoo! News

Pros:
- Quick understandability
- Saving display space

- Biggest news portal in Japan
  - PV/month: 15,000,000,000+
  - Editors' choice feature ->



News articles delivered by providers

1. Pick up important news articles

2. Put a new shorter headline, called **short title**

Professional editors

Editor's choice feature

Short title generation as editing support

- Purpose: To generate short title candidates to help editors
- Task: Translation from (headline, lead) to short title
  - Lead is a short version (summary) of the article

Selected news article

List of news articles

Example of (short title, headline, lead)

Phrase order is changed

	Japanese	English translation
Short title	首相忖度ないと言い切れず	The prime minister cannot say that there is no surmise
Headline	忖度なかったと言い切ることはできない＝加計問題で安倍首相	It cannot be said that there is no "sontaku (surmise)" with absolute certainty. The prime minister Abe said about the problem of "Kake Gakuen (Kake school)".
Lead	安倍晋三首相は１４日午後行われた参院予算委員会の集中審議で、加計疑惑などを巡り、官僚側から首相に対する忖度（そんたく）があったのではとの指摘に対して「忖度があったかどうか、忖度される側には分かりにくい面もある」と述べた。「忖度がなかったと言い切ることはできない」としつつ、「ごまをするための忖度は求めていない」などと説明した。塚田一郎委員（自民）への答弁。	Prime Minister Shinzo Abe said, in an intensive deliberation with the House of Councilors Budget Committee held on the afternoon of the 14th, as an answer to a question about whether bureaucrats surmised to the prime minister regarding the Kake suspicion, "It is difficult to understand whether there is a sontaku (surmise)". He said "It cannot be said that there was nothing wrong," while explaining that "I do not need to be obsequious". An answer to Ichiro Tsukada (LDP).

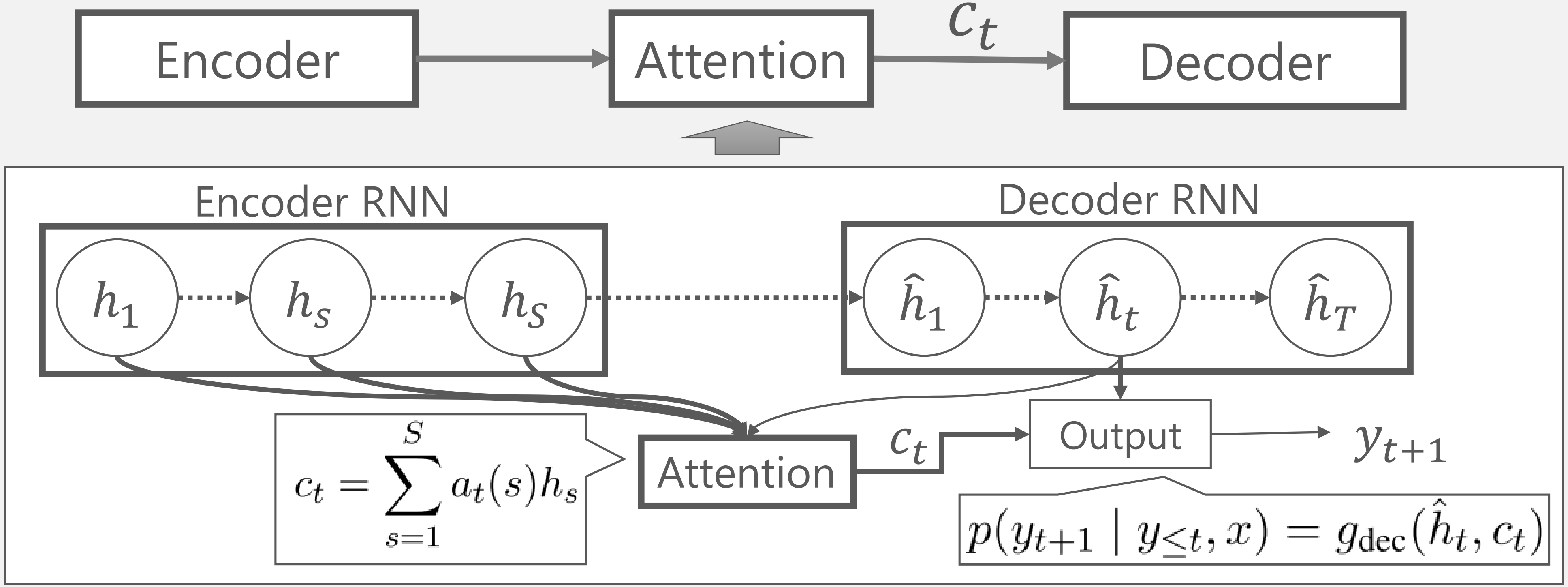
Lengths are different

Short title generation task is not so easy

YAHOO! JAPAN

Encoder-decoder model with attention

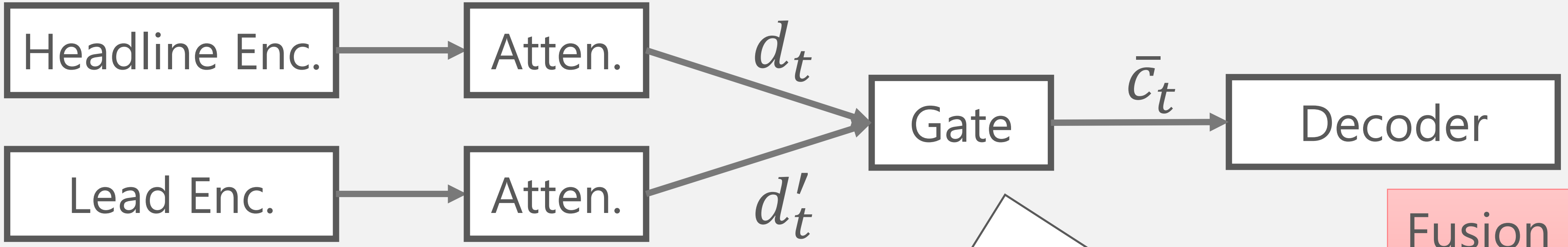
- Conditional language model consisting of two RNNs
  - Described by three components (encoder, attention, decoder)



Attention calculates a context c_t from the encoder's states h_s

Proposed method: GateFusion

- Combine headline and lead contexts w/ gating mechanism



Existing work (Hori+ 2017) used an attention mechanism

$$\bar{c}_t = \alpha d_t + \beta d'_t$$

Fusion based on scalar weights

Gating mechanism:

$$w_t = \sigma(W[d_t; d'_t; \hat{h}_t]),$$

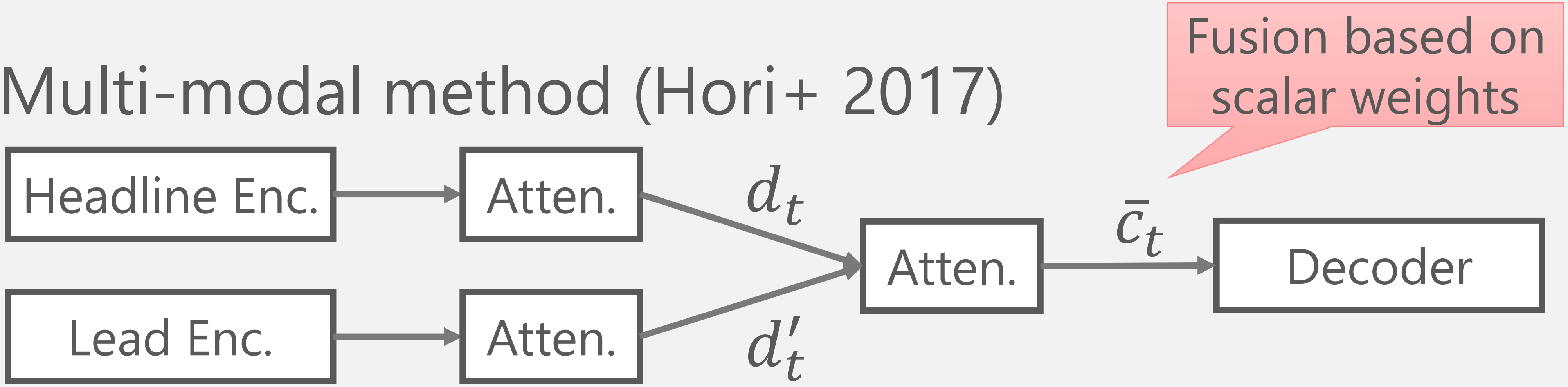
$$w'_t = \sigma(W'[d_t; d'_t; \hat{h}_t]),$$

$$\bar{c}_t = w_t \odot d_t + w'_t \odot d'_t,$$

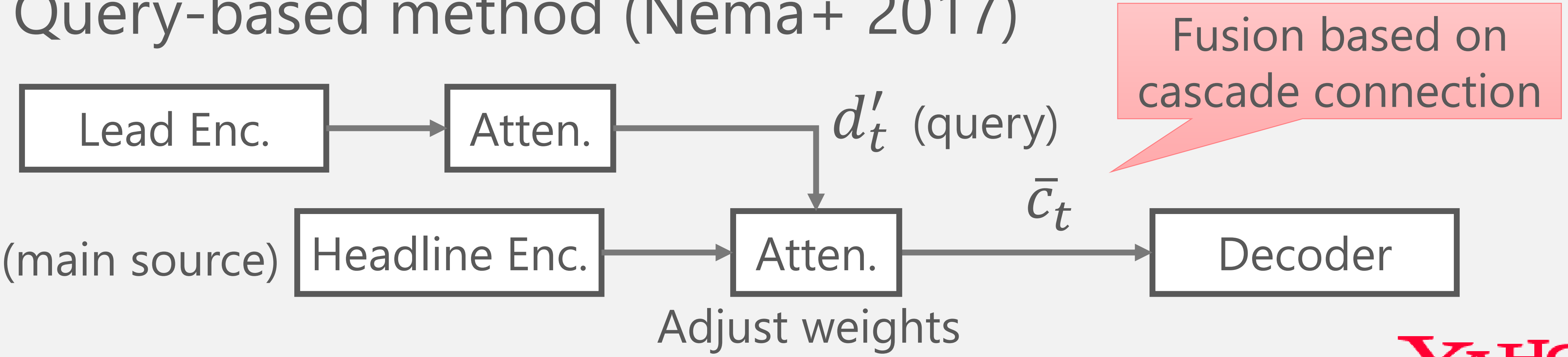
Fusion based on vector weights

Baselines with multiple encoders

- Multi-modal method (Hori+ 2017)



- Query-based method (Nema+ 2017)



Training dataset

- 263K triples of (headline, lead, short title) in Yahoo! News
  - Training (90%), validation (5%), testing (5%)
- Statistics:

	Headline	Lead	Short title
Average length	24.87	128.49	13.05
Character type	3618	4226	3156

  - Extractively solvable instances: 20%
    - Characters in each short title are completely covered by the headline
  - Edit distance of headlines and short titles: 23.74
    - Short titles cannot be easily created only from headlines

Model and training settings

- Implemented on OpenNMT
- Headline encoder: BiLSTM
- Lead encoder: CNN (Kim, 2014)
  - To reduce the computational time
- Ensemble of 10 models
- Hyper-parameter settings are listed in the right table

Hyper-parameter	Value
# of layers (RNN, CNN)	3
# of units (embedding)	200
# of units (RNN, CNN)	400
# of units (context)	400
Window width of CNN	7
Dropout rate	0.3
Learning rate	0.05
Momentum rate	0.8
Learning_decay rate	0.85
# of epochs	20
Batch size	64
Beam width	5

Human evaluation by crowdsourcing

- Two crowdsourcing tasks for readability and usefulness
  - Average score of 10 workers for each of 1,000 outputs
- Readability (four-point scale)
  - How readable a short title was
- Usefulness (four-point scale)
  - How useful a short title was compared to the headline

Evaluation results (1/2)

- Our model performed well for the usefulness measure

		Readablity	Usefulness	Average =(R+U)/2
Correct titles	Editor	3.62	3.18	3.40
First 13 chars	Prefix	2.72	2.38	2.55
Single enc.	OpenNMT	3.53	3.16	3.35
Multi enc.	MultiModal	3.51	3.12	3.32
	QueryBased	3.52	3.11	3.32
Our model	GateFusion	3.50	†**3.22**	3.36

Complicated expressions

Aggressively copy characters

Evaluation results (2/2)

- Our model performed well for the usefulness measure

		Readablity	Usefulness	Average =(R+U)/2
Correct titles	Editor	3.62	3.18	3.40
First 13 chars	Prefix	2.72	2.38	2.55
Single enc.	OpenNMT	3.53	3.16	3.35
Multi enc.	MultiModal	3.51	3.12	3.32
	QueryBased	3.52	3.11	3.32
Our models	GateFusion	3.50	†**3.22**	3.36
Gate+Query	HybridFusion	†**3.55**	†**3.22**	†**3.39**

Close to Editor

QueryBased helped GateFusion generate headline-style outputs

	Input and generated title (Japanese)
Headline	逆境をチャンスに変えたダルビッシュの進化
Lead	レンジャーズのダルビッシュ有（29）が 28 日、本拠地で行われたパイレーツ戦で [...]
Editor	術前より進化 ダルの肉体改造
OpenNMT	逆境をチャンスに変えた進化
HybridFusion	ダル 逆境をチャンスに変えた

Last word tends to be important

Best baseline

Our best model

	English translation
Headline	Evolution of Darvish; turning adversity into opportunity.
Lead	Yu Darvish (29) in Rangers took a mound for the first time in 1 year and 9 months with Pirates [...]
Editor	Dar sculpted his body better than before surgery.
OpenNMT	Evolution; turning adversity into opportunity.
HybridFusion	Dar turned adversity into opportunity.

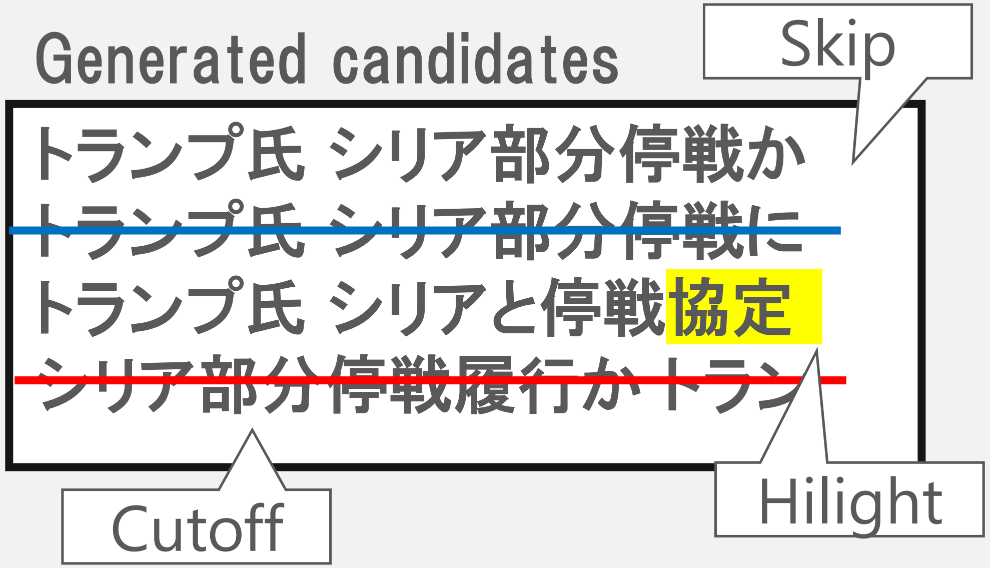
Our model worked even in this real-world application

Editing support tool

- Editors can check candidates when creating short titles

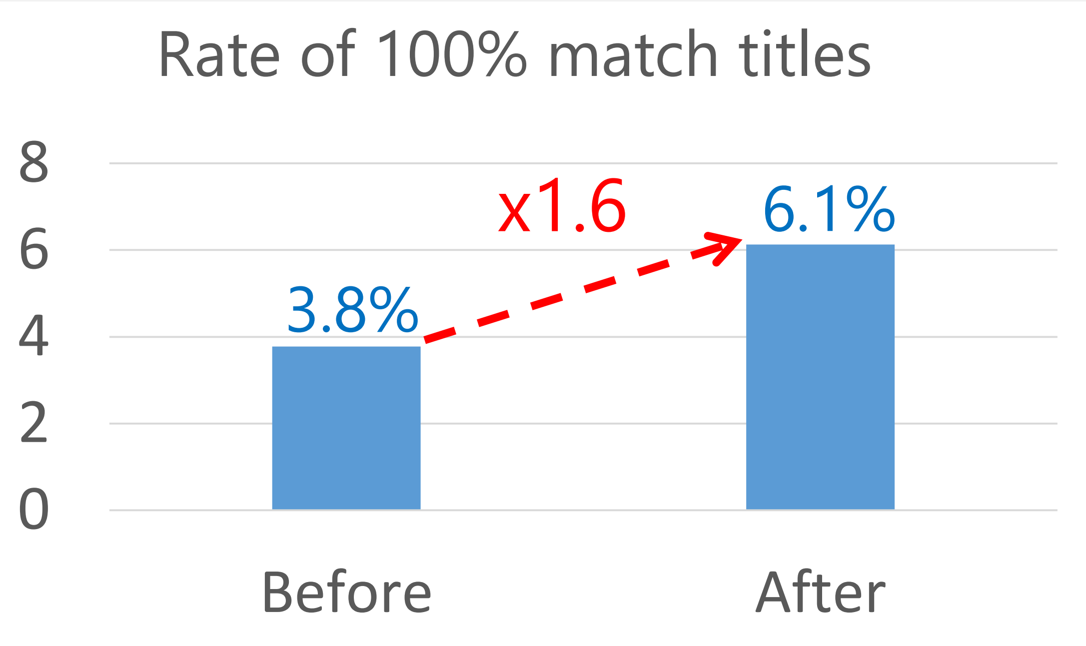
Functionalities in the tool

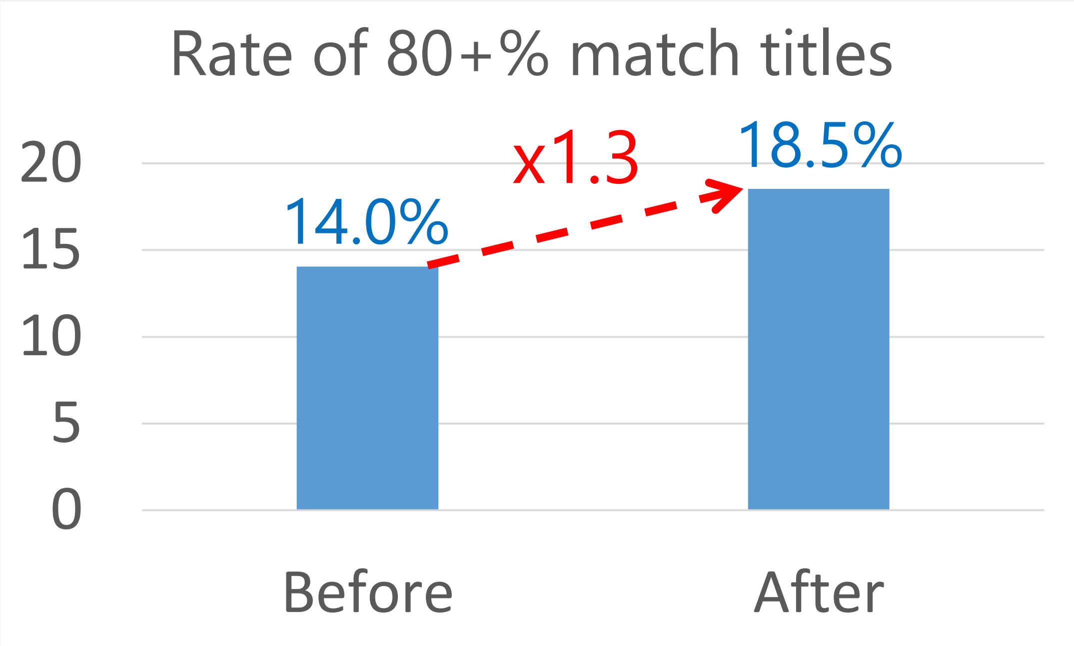
- Cutoff unpromising candidates
  - If perplexity>x
  - To keep the system quality
- Skipping redundant candidates
  - If edit distance<y
  - To display various outputs
- Highlighting unknown characters
  - If not in the article
  - To encourage fact checking



Effect of the tool release

- Editors’ behavior in three weeks before/after the release
  - Rate at which an editor’s title matches the generated one by X%





Editors began to refer to generated outputs after the release

Conclusion

- Short titles were successfully generated for editing support
- Editors began to refer to generated titles of our system

- Future work
  - Verify how much our model can affect click-through rate
    - Need a much safer model to avoid generating fake titles

- Acknowledgements
  - We would like to thank editors and engineers in the news service who continuously supported our experiments

Thank you for your attention!