# Extractive Headline Generation Based on Learning to Rank on Community Question Answering

Tatsuru Higurashi[1]  Hayato Kobayashi[1] [2] Takeshi Masuyama[1] Kazuma Murao[1]

[1]Yahoo Japan Corporation    [2] RIKEN AIP

# Contents

- **Motivation and Approach**

- **Proposed method**

- **Experiment and Results**

- **Conclusion**

# Contents

- **Motivation and Approach**


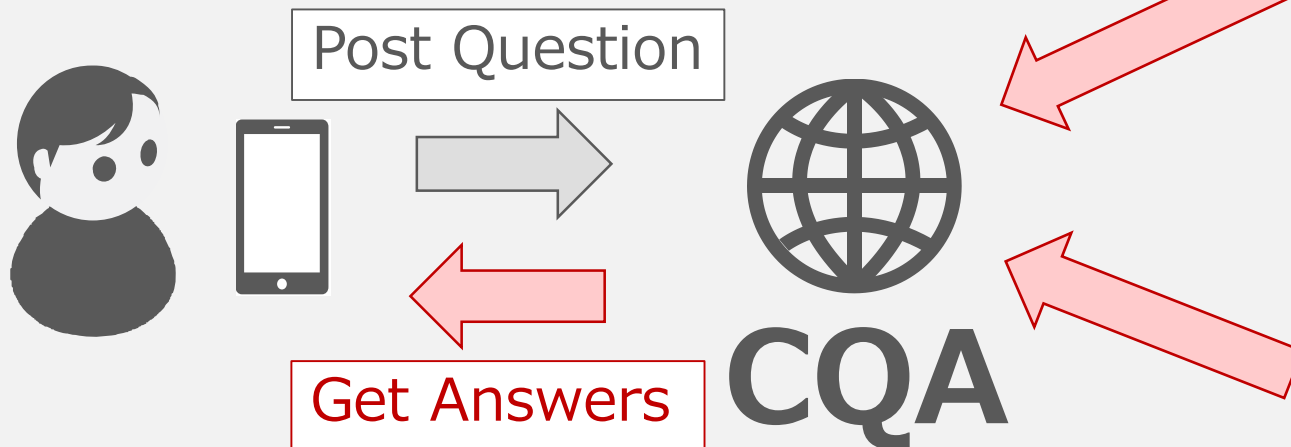- Proposed method


- Experiment and Results


- Conclusion

**Question**

Advice please.
A dog kept in the next house barks from morning to night.
…
How can I effectively manage this problem?

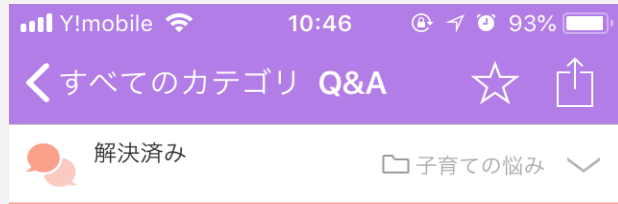**Answer of User A**
No solution other than moving.

**Answer of User B**
Please contact the public health center.

Post Question

Get Answers

CQA

# Push Notification in CQA

- **Push Notification obtain quick answers**
  Directly linked to the quality of CQA

Push Notification of Question

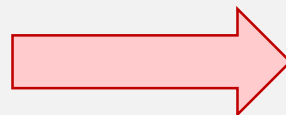すべてのカテゴリ **Q&A**

解決済み　　　　　子育ての悩み

**Contents of Posted Question**

はじめまして、よろしくお願いします、30
代の男性です。
アドバイスお願いします。

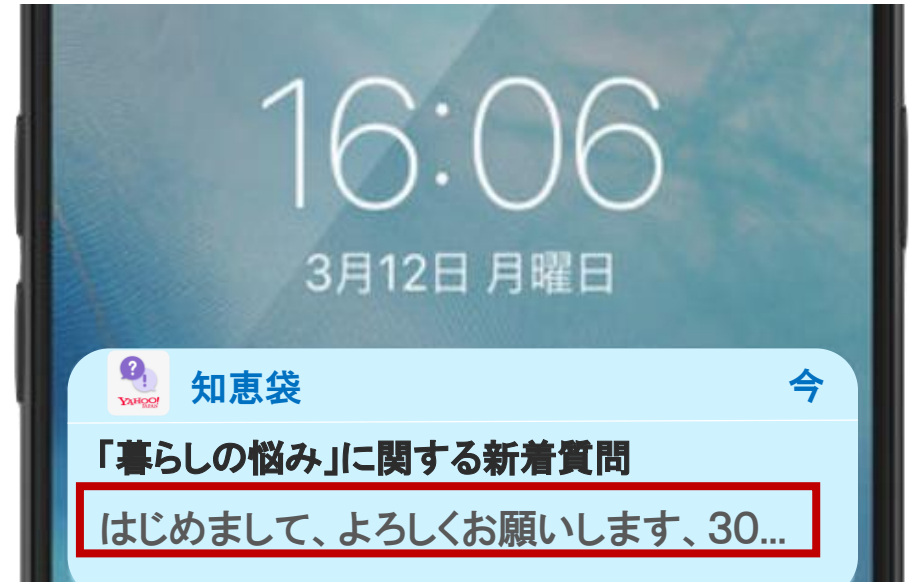隣の家が飼っている犬が朝から晩まで
吠え続けます。
近所の人たちも飼い主に注意を促して
いるのですが、改善が見られませ

**Push
Notification**

**CQA
Yahoo! Chiebukuro**

知恵袋　　　　　　　　今

「暮らしの悩み」に関する新着質問

はじめまして、よろしくお願いします、30…

**Headline of Posted Question**

**Respondent**

# Snippet Extraction Makes Headline Informative

- ## Approach
  Extract a mid-substring as a snippet

**Posted Question**

Nice to meet you, thank you in advance. I'm a man in my thirties.
Advice please.

A dog kept in the next house barks from morning to night.
Neighbors have given the owner cautions against it,

隣の家が飼っている犬が朝から晩まで吼え続けます。
近所の人たちも飼い主に注意を促して

**Prefix Headline:** Uninformative

Nice to meet you, thank you in advance. I'm a...

**Snippet headline:** Informative

...A dog kept in the next house barks from morning...

# Contributions

1. Show empirical evidence that snippet headlines are more effective than prefix headlines

2. Propose extractive headline generation method based on learning to rank

3. Create Japanese dataset including headline candidates with "headline-ness" scores by crowdsourcing

# Advantages of Snippet Headlines

- ## Effective

**Average Answer Rate**

×2.4

0.31%
**Prefix Headline**

0.75%
**Snippet Headline**
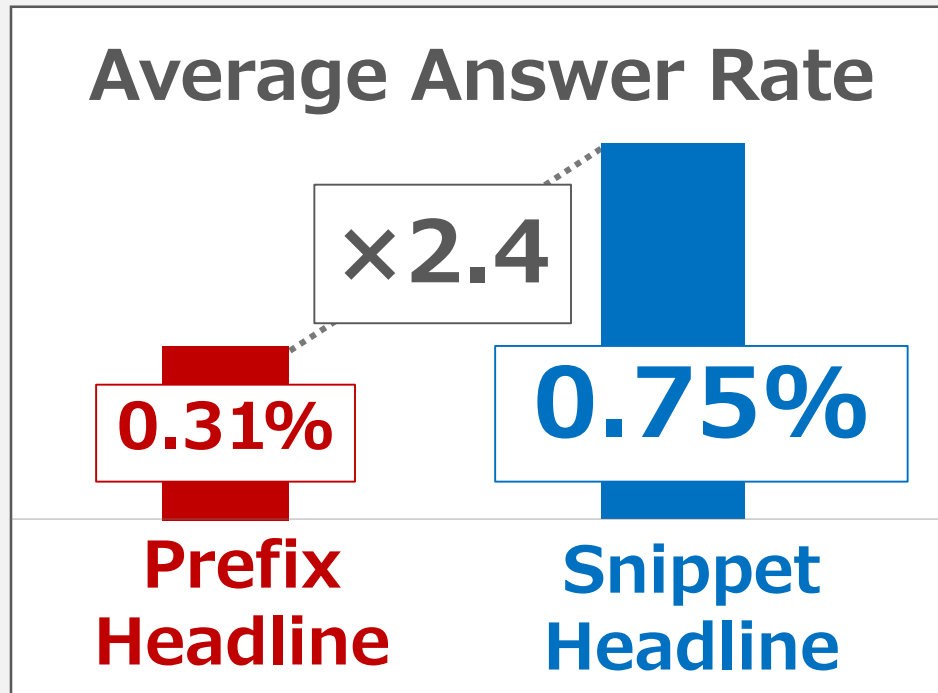
A/B testing on Yahoo! Chiebukuro push notifications of smartphones

- ## Practical

**Snippet headlines never include generative errors**

- Commercial services (Yahoo! Chiebukuro) cannot accept generative errors

- Abstractive methods cannot completely avoid generative errors

- **Neural Headline Generation:**

  - [Takase+ '16][Chopra+ '16][Kiyono+ '17][Zhou+ '17][Raffael+ '17]
- **Sentence Compression:**

  - [Woodsend+ '10][Alfonseca+ '13][Colmenares+ '15]
- **Extractive Summarization:**

  - [Kobayashi+ '15][Yogatama+ '15][Cheng&Lapata '16][Nallapati+ '17]
- **CQA Data:**

  - [Bhasker+ '13] [Ishigaki+ '17][Surdenau+ '08][Celikyilmaz '09]

Our research is first attempt to address extractive headline generation for CQA service with substring of question based on learning to rank

# Contents

- **Motivation and Approach**


- **Proposed Method**


- **Experiment and Results**


- **Conclusion**

# Overview of Our Proposed Method

**Posted Question**
Nice to meet you, thank you in advance. I'm a man in my thirties. Advice please.
A dog kept in the next house barks from morning to night...
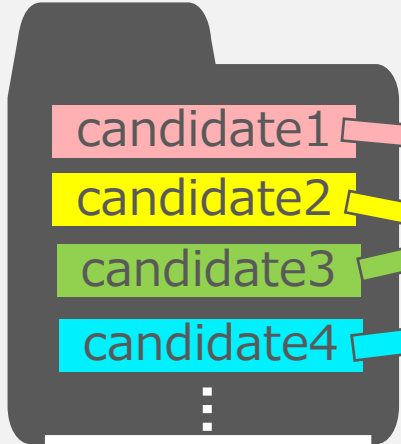
**Ranked Headline**
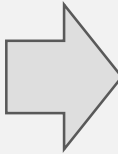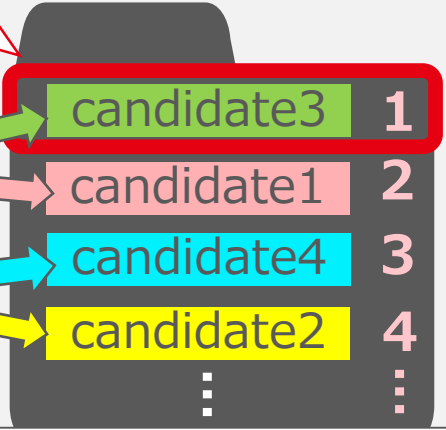...A dog kept in the next house barks from morning to night...

Question

Candidate Generation

candidate1
candidate2
candidate3
candidate4
⋮

Candidates

Candidate Ranking

candidate3  1
candidate1  2
candidate4  3
candidate2  4
⋮

Ranked Headline

# Candidate Generation

**Generated Candidate Snippet**

はじめまして、よろしくお願いします、私 …
*Nice to meet you, thank you in advance, I'm a...*

Cut subsequent sentences if over 20 Japanese characters

Ellipsis … アドバイスお願いします。隣の家が飼っ … Ellipsis
*… Advice please. A dog kept in the next house …*

… 隣の家が勝っている犬が朝から晩まで吼 …
*… A dog kept in the next house barks from morning …*

- Make sentence which starts from beginning of each sentences of question.

- Cut subsequent sentences if it has over 20 Japanese characters.

- Put ellipsis at front and end of substring.

- **Pairwise Learning to Rank**

  - L2-regularized L2-loss linear rankSVM

$$\min_{\omega} \quad \frac{1}{2}\omega^{\top}\omega + \mathcal{C}\sum_{(i,j)\in P}\ell\left(\omega^{\top}\widetilde{x}_i - \omega^{\top}\widetilde{x}_j\right) \qquad (\ell(d) = \max(0, 1-d)^2)$$

$$P = \{(i,j)\,|\,q_i = q_j, y_i > y_j, (x_i, y_i, q_i) \in D, (x_j, y_j, q_j) \in D\}$$

$x$: *Candidate headline*
$y$: *Vote score*
$q$: *Question*

[Lee 2014] Lee, Ching—Pei and Lin, Chih—Jen: Large—scale Linear Ranksvm. Neural Computation, 26(4)

# Data Creation: Crowdsourcing

## Instruction for workers

Select the best option from the list so that users can guess the content of the question and distinguish it from other ones.

| Randomly Sorted Headline Candidate | score |
|---|---|
| 1 … Neighbors have given the owner cautions against … | 2 |
| 2 … This area has only private houses, not rented … | 0 |
| 3 … How can I effectively manage this problem? | 1 |
| 4 … However, I will go crazy if I have to keep enduring … | 0 |
| 5 … A dog kept in the next house barks from morning … | 6 |
| 6 Nice to meet you, thank you in advance, I'm a … | 0 |
| 7 … Advice please. A dog kept in the next house … | 1 |

Number of votes by
10 workers per question

# Crowdsourcing Results

- Ratio of questions whose prefix headlines were most voted candidates was **61.8%**

= Room for improvement for prefix headline was up to **38.2%.**
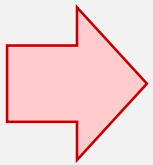
➡ **Our Goal**
- Improve uninformative headlines of 38.2%
- Remaining 61.8% unchanged

# Contents

- **Motivation and Approach**

- **Proposed Method**

- **Experiment and Results**

- **Conclusion**

# Features for Ranking Model

- **Features**

  - Bag-of-Words: 30,820 dimension sparse vector based on tf-idf

  - Embedding: 100 dimension dense vector based on doc2vec

  - Position: 10 dimension binary vector representing candidate position

# Compared Methods

- **Simple Baselines**

  - *Prefix:*  Select first candidate
  - *DictDel:* Delete uninformative sentence with rule (Used in A/B testing)
  - *Random:* Select randomly

- **Unsupervised Baselines**

  - *ImpTfidf:* Select most important candidate with highest tf-idf value
  - *SimTfidf:* Select most similar candidate to original question with cosine similarity
  - *LexRank:* Select candidate with highest score based on LexRank (Erkan&Radev 2004)

- **Supervised Baselines**

  - *SVM:* Select candidate with highest confidence learned as classification task
  - *SVR:* Select candidate with highest predicted votes learned as regression task

# Evaluation Metrics

- ## Average Votes

  - Measures how appropriate candidates selected by each method
  - Determines the overall performance of each method
  - Corresponds to DCG@1

$$\text{Average votes} = \frac{\text{Sum of votes for the best candidates for all questions}}{\text{No. of questions}}.$$

- ## Change Rate

  - Measures how much each method changed the default prefix headline
  - Determines the impact of application to actual CQA service.

$$\text{Change rate} = \frac{\text{No. of questions where the best candidate is not the prefix headline}}{\text{No. of all questions}}.$$

# Results: Quantitative Analysis

| Method | Avarage Votes | Change Rate |
|---|---|---|
| *Ideal* | 5.56 | 38.2 |
| *MLRank(ours)* | ① 4.28 | 9.9 |
| *Prefix* | ② 4.19 | 0 |
| *DictDel* | ③ 4.23 | ④ 2.2 |
| *Random* | 1.39 | 85.9 |
| *ImpTfidf* | 1.68 | ⑤ 81.1 |
| *SimTfidf* | 2.27 | 79.3 |
| *LexRank* | 2.27 | 55.7 |
| *SVM* | 4.09 | 16.7 |
| *SVR* | 3.00 | 52.5 |

① *MLRank(ours)* performed the best among all methods.

② *Prefix*(First sentence) can be a good summary.

③ *DictDel*(Rule-Based) was more useful than Prefix.

④ Change rate of *DictDel* was small, which means small impact on service.

⑤ Change rates of unsupervised methods were high, but the overall performances were low

# Results: Qualitative Analysis

- **Examples of prefix headline and snippet headline**

| Prefix Headline | Snippet Headline |
|---|---|
| **カテ違いならゴメンナサイ**。今、財布が破… | …今、財布が破**れてツライ状況です。新し**… |
| **I am sorry if the category is wrong.** Now, my wallet is torn… | … Now, my wallet is torn, **and I'm having a hard time. A new one** … |
| **27歳女です。**環境的になかなか新しい出… | …環境的になかなか新しい出**会いがなく**… |
| **I am a 27-year-old woman.** Owing to my environment, there is little chance of new … | … Owing to my environment, there is little chance of new **encounters with men** … |

Uninformative expressions are successfully excluded, and informative experssions are added

# Contents

- **Motivation and Approach**

- **Proposed Method**

- **Experiment and Results**

- **Conclusion**

# Conclusion

## Addressed a snippet headline generation task for push notifications of CQA

- Showed empirical evidence that snippet headlines are more effective than prefix headlines 2.4 times in average answer rate

- Proposed extractive headline generation method based on learning to rank

- Created dataset including headline candidates with "headline-ness" scores by crowdsourcing

## Future Work

- Investigate effectiveness in practical situations on web service.

- Make the dataset publicly available.

# Thank you

YAHOO! JAPAN