Human Values Estimation on News Articles through BERT-extracted Opinion Expressions

Yihong Han[1]、Yoko Nishihara[2]、Junjie SHAN[3]

[1] Graduate School of Information Science and Engineering, Ritsumeikan University
[2] College of Information Science and Engineering, Ritsumeikan University
[3] Ritsumeikan Global Innovation Research Organization, Ritsumeikan University

Background

- Web 2.0 era: everyone can post their opinion on the Internet.

- There are so many opinions from various viewpoint in the cyber world.

- The continuous development of "information optimization" and "recommendation system".

- The "echo chamber": the situations in which beliefs are amplified or reinforced by communication and repetition inside a closed system and insulated from rebuttal.

- Influence of "echo chamber": misunderstanding to the real world, thought fragmentation in society, obstacle from consensus building

Concept

- Reason of "echo chamber": the limited access to the opinions and thoughts

- Solution: show the various opinions based on different human values to people

- Necessary tech: human values estimation from article

- Hypothesis: human values are primarily included in the sentences which reflect the writer's opinion.
  "Kyoto has been the capital of Japan for over a thousand years"→Fact
  "I love Kyoto's historical atmosphere."→Opinion
  "Kyoto's landscape as the ancient capital of Japan should be carefully preserved."→Opinion with human values

Proposed Method

- [Input]　News articles is the input to the proposed system.

- [Step 1]　Opinion sentences extraction & Opinion sentences set building

- [Step 2]　Human values estimation for each opinion sentence

- [Output]　Human values distribution of the whole article
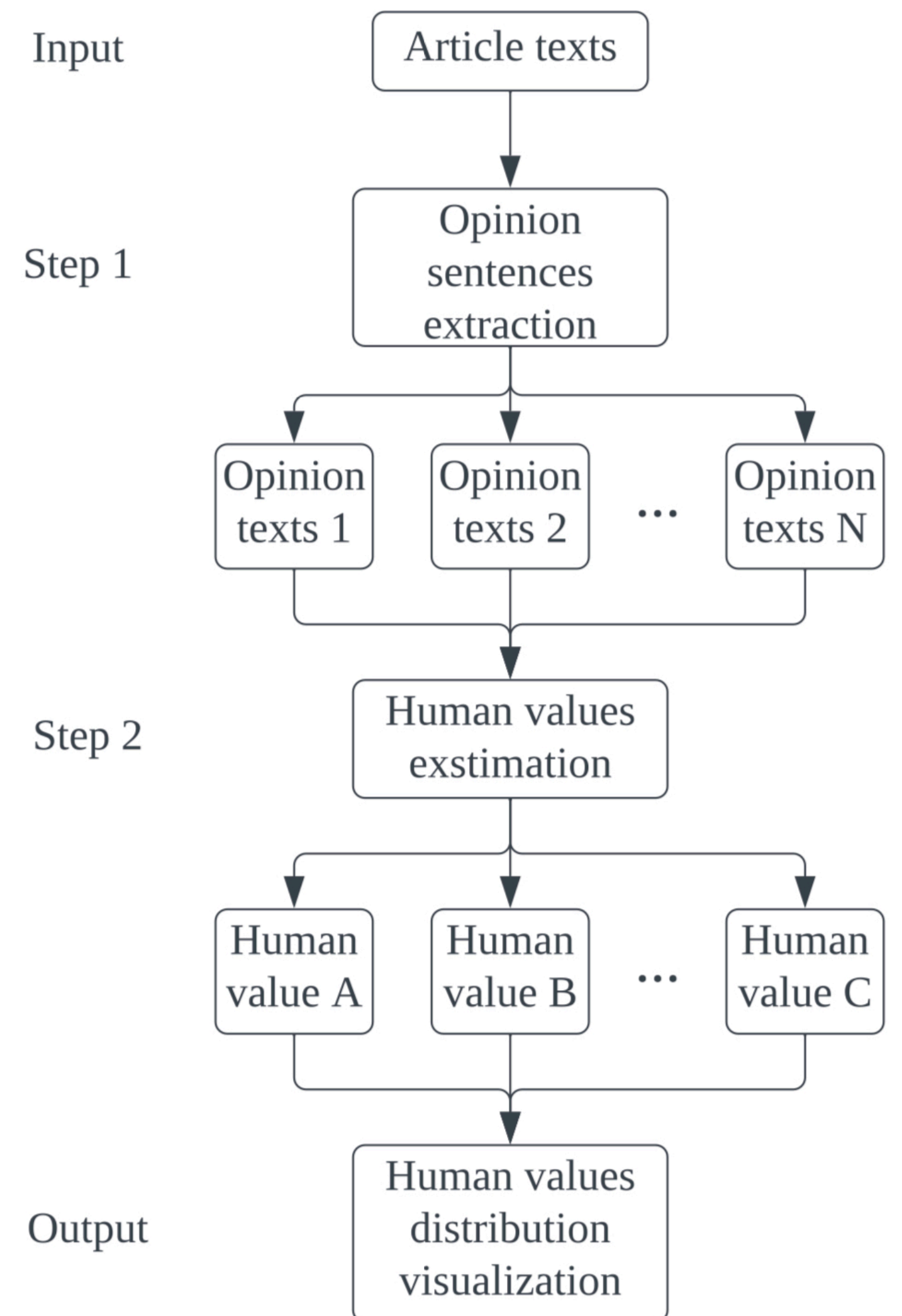
Opinion Extraction

- Categories of opinion sentences:
  "Needing to be done or requiring a response", "Demand", "Emotion", "Critique", "Merit", "Adoption" and "Event"
  Red items are considered to combine human values, blue items are not.

- Training data:
  23 editorial articles from Mainichi Shimbun(毎日新聞) article dataset

  Each article is split to sentences, every sentence is categorized into "opinion sentences" and "non-opinion sentences".
  There are 224 opinion sentences and 458 non-opinion sentences in the training dataset.

- An example of opinion sentence: "I hope that all government agencies will work closely together and take all possible measures to respond to the situation."
  An example of non-opinion sentence: "This is the largest earthquake ever recorded in Japan."

- Training result: accuracy reached 96%

Human Values Estimation

- Training data:
  211 editorial articles from Mainichi Shimbun(毎日新聞) article

  dataset (every article has a human values label)
  Extract opinion sentences with the opinion extraction module
  described in last page
  Article's label→sentence label

- Categories of human values:
  Human welfare, Effectiveness, Importance, Power, Law and
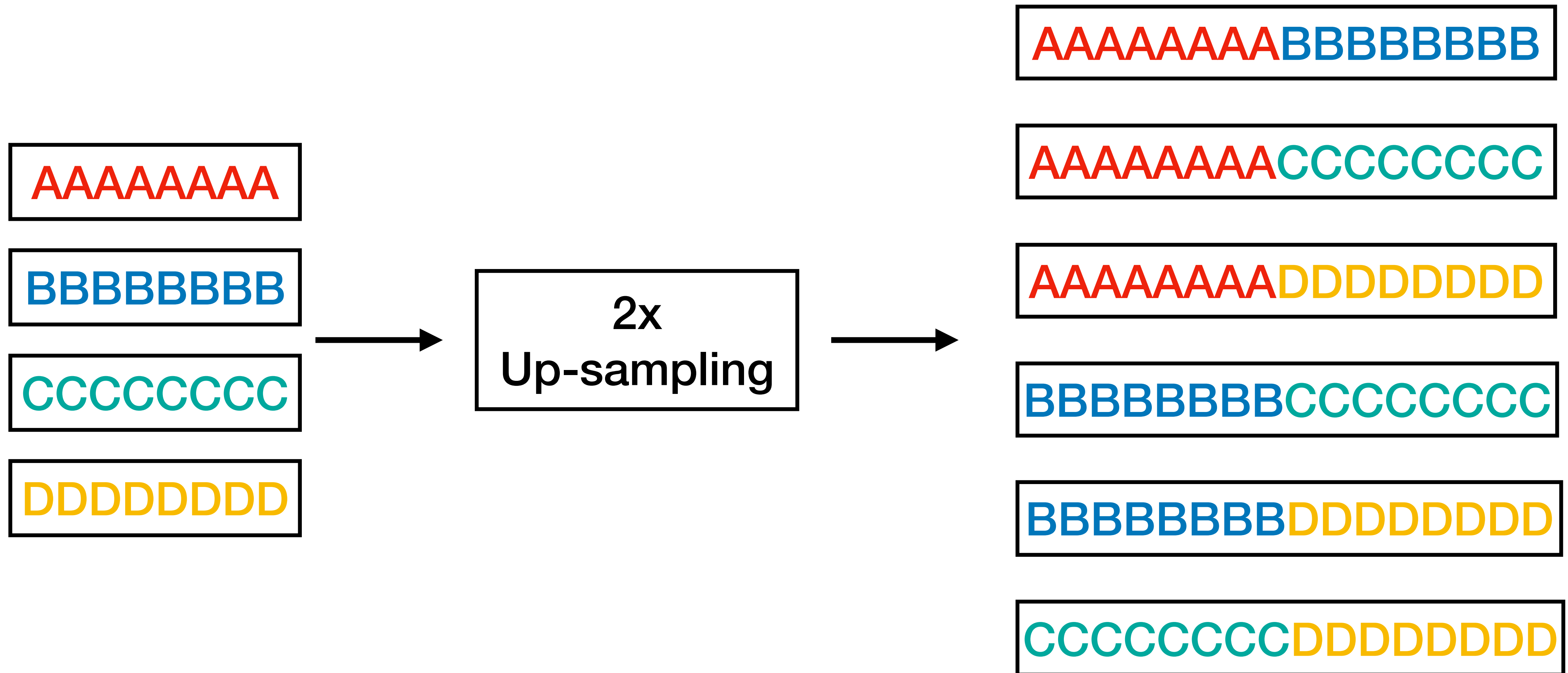 order

- Training result:
  accuracy reached 40%

Categories	Article numbers
Human welfare	57
Effectiveness	56
Importance	48
Power	33
Law and order	17

Categories	Sentence numbers
Human welfare	514
Effectiveness	594
Importance	449
Power	375
Law and order	174

Data Up-sampling

- It is difficult to estimate human values from only one sentence.

- We increase the information per input by using more texts to build the training dataset.

- Multiple sentences from an article are concatenated to build one input.

- The number of variations when choosing r sentences from N sentences: $_NC_r$

- The purpose of up-sampling:
  information per input↑　number of input↑　variations of training dataset↑

- An example:
  Before up-sampling→sentence number:10, 1 sentence per input, number of input: 10
  After up-sampling→sentence number:10, 2 sentence per input, number of input: 45 ($_{10}C_2$)

The Image of Data Up-sampling

AAAAAAAA

BBBBBBBB

CCCCCCCC

DDDDDDDD

→ 2x Up-sampling →

AAAAAAAABBBBBBBB

AAAAAAAACCCCCCCC

AAAAAAAADDDDDDDD

BBBBBBBBCCCCCCCC

BBBBBBBBDDDDDDDD

CCCCCCCCDDDDDDDD

Data Up-sampling

- 1 sentence per input→2 sentences per input→3 sentences per input

- An example:
  "The entire nation must come together and do everything in its ability to provide a help."
  ↓
  "The entire nation must come together and do everything in its ability to provide a help. The entire nation must work together to overcome this disaster."

Categories	Input number before	Input number after
Human welfare	514	2670
Effectiveness	594	3598
Importance	449	2357
Power	375	2431
Law and order	174	887
Total	2106	11943

Experiment: Opinion Extraction

- Dataset building for experiment:
  17 articles (different from training dataset) chosen, split as sentences, annotated as "opinion" and "non-opinion"
  165 opinion sentences and 244 non-opinion sentences in the experiment dataset

- Experiment result:

	Predicted as Opinion	Predicted as Non-opinion
Annotated as Opinion	144	21
Annotated as Non-opinion	28	244

- Precicsion 83.7%, Recall 87.3%, F1-score 85.5%

Discussion: Opinion Extraction

- The features of opinion expressions can be recognized with a small dataset training.

- The main reason: grammatical characteristic of opinion expressions (a Japanese language feature)
  "needing to be done or requiring a response"→some verb + べき (means "should
  do something")
  "demand"→some verb + てほしい (means "want or wish for some action")

- News articles is a formal article style: grammatical stability
  There are regular uses of vocabulary or phrases in this field.
  Editors have limited choice to represent a conception or meaning.

- The performance may change for spoken language use.

Experiment: Effectiveness of Up-sampling

- Data up-sampling was conducted in both training dataset building and test dataset generating.

- The purpose of up-sampling in training dataset building:
  information per input↑　number of input↑　variations of training dataset↑

- The purpose of up-sampling in test dataset generating: information per input↑

- Experiment results:

Accuracy　　　　　Extension Level Input Level	1	2	3
1	40.5% (case 1)	97.0% (case 4)	90.5% (case 5)
2	-	98.6% (case 2)	99.1% (case 6)
3	-	-	99.7% (case 3)

Discussion: Effectiveness of Up-sampling

- Compared to the original case, the accuracy rose from 40% to over 90%.

- There isn't a huge difference between two-sentences case and three-sentences case.
  When the input has over two sentences, the model achieves good performance.

- Multiple sentences training increased the information of training dataset and obtained a better performance.

- A closed corpus have similar literary pattern in expressing opinions and human values between training dataset and test dataset.

- Multiple sentences input also helps prediction accuracy (90.5%→99.7%).

Human Values Estimation Examples

Example Sentences	Estimation Results
As long as the effects of low-dose radiation on the human body are unclear, it goes without saying that careful decontamination is necessary.	Human welfare
It is an important mission of the press to interview the people involved, identify problems from multiple perspectives, and present lessons learned.	Importance
The government is reviewing the structure of nuclear safety regulation, including its separation from the Ministry of Economy, Trade, and Industry (METI), but the problem will not be solved unless the structure that creates entanglement is addressed.	Law and order
No nuclear reactors will be restarted unless we can gain the understanding of the local residents.	Human welfare
We would like to request the LDP and other opposition parties to participate constructively in the debate. This is the responsibility of the legislature.	Power

Summary

- This paper proposed a human values estimation method with an opinion extraction approach.

- The proposed method extracts the opinion sentences from the texts and estimates the human values included in the opinion sentences.

- We proposed a data up-sampling method to increase the information of training dataset and test input.

- The data up-sampling method concatenates multiple sentences into one.

- The accuracy of opinion sentence extraction was 92%.
  In the evaluation test, the model reached a F1 score of 85.5%.

- After data up-sampling method deployed, the accuracy raised from 40.5% to 99.7%.
  The results showed that the human values of opinion sentences could be estimated with high accuracy.

- We plan to utilize the proposed method to estimate the human values of SNS posts and comments on news articles.

Thank you for your listening