WER99 at the NTCIR-15 QA Lab-PoliInfo-2 Classification Task

Yuichi Sasazawa
Tokyo Institute of Technology
yuichi.sasazawa[at]nlp.c.titech.ac.jp

Naoaki Okazaki
Tokyo Institute of Technology
okazaki[at]c.titech.ac.jp

ABSTRACT

In this paper, we report our method for the stance classification task of NTCIR15 QA Lab-PoliInfo-2. There are two types of stances that we need to detect: stances that are explicitly stated and those that are not stated in utterances. We designed a set of rules to recognize an explicit mention of a stance for a bill. When a party does not explicitly mention a stance, our method uses clues in the bill name to predict a stance. The method achieved the highest performance (99.75% accuracy) among the participants on the test data.

TEAM NAME

wer99

SUBTASKS

Stance Classification (Japanese)

1 INTRODUCTION

In this paper, we report our method for the stance classification task of NTCIR15 QA Lab-PoliInfo-2 [2]. In this task, we need to predict the stance of a political party on each bill. Participants use the minutes of the Tokyo Metropolitan Assembly meetings as input. The minutes of the meetings include various pieces of information such as utterances of politicians, reports, and a list of bills to be discussed. There are two types of stances that we need to detect: (1) stances that are explicitly stated and (2) those that are not explicitly stated in the utterances. In this study, we provide different approaches for detecting these stances accurately.

When politicians explicitly state their party's stance in the utterance, we apply a rule-based algorithm to predict whether the party is either for or against the bill. As used in the parliament, utterances are so formal that we can detect stances rather easily. When politicians do not state their party's stance in the utterance, we predict the stance by analyzing the bill names. In addition, we used several methods to improve the accuracy of stance detection.

We achieved a 99.75% accuracy in the automated evaluation metric. This was the highest performance of all participants. The result of the human evaluation metric also showed that our method was much more accurate (98.2% accuracy) than those of the other methods. These results demonstrated that our method is very effective for this task.

The rest of the paper is organized as follows. Section 2.1 describes the rules for detecting stances from utterances. Section 2.2 presents the methods for predicting a stance from a bill name. Section 2.3 reports other useful techniques for stance detection. Finally, Section 2.4 provides an overview of the overall classification process.

Throughout this paper, we present examples written in Japanese with English translations.

2 METHODS

2.1 Detecting stances from utterances

2.1.1 Plenary session. Major parties such as "自民党 (the Liberal Democratic Party)," "民主党 (the Democratic Party)," and "公明党 (the Komeito)" explicitly state their stances on bills that they are opposed to in the plenary session. Here is an example.

> "私は、都議会公明党を代表して、共産党による議員提出議案第一号に反対し、平成十八年度東京都一般会計予算ほか、知事提出全議案に賛成する立場から討論を行います。(On behalf of the Komeito Party, I am going to debate from the standpoint of opposing bill No. 1 submitted by the Communist Party and approving all bills submitted by the governor, including the fiscal year Heisei 18 Tokyo Metropolitan Government's General Account Budget.)"

We can easily extract the bills that the politician opposes by using a rule-based method. Figure 1 shows an overview of the method.

The method first identifies sentences wherein the politicians declare their party's stance on a bill. More specifically, we extract sentences that include words such as "賛成 (agree)" and "反対 (oppose)" in the minutes. We then identify the party that declares its stance. In most cases, politicians declare their party's stance in the following format: "xxx を代表して (On behalf of the xxx)," where xxx is the party name. We refer to the list of parties that we have prepared in advance and identify the party from this pattern.

Next, we divide the utterance into two segments: the segment where the politicians declare the bills that their party opposes and that where they declare the bills that their party agrees with. In most cases, the politicians declare their party's stance in the following format: "xxx を代表して、... に反対、... に賛成、... に反対します。(On behalf of the xxx, we are opposed to ..., agree with ..., are opposed to, etc.). " Therefore, we divide the sentence into segments by using the words "賛成 (agree)" and "反対 (oppose)" as clue words.

Finally, we identify the bills that a party is opposed to. We consider two cases. In the first case, a bill name is specified by a bill number, e.g., "第二十号議案 (Bill No. 20)." In this case, we extract the phrase "xxx 号 (where xxx is a Chinese numeral)" and identify the corresponding bill. In the second case, a bill name is directly declared. In this case, we refer to the pre-prepared list of bill names and extract the bill that a party opposes.

A number of researchers have used machine learning (ML)-based methods for sentence/stance classification [1, 3, 5]. Although we explored a deep neural network-based method, the rule-based method was sufficient and accurate to detect stances.

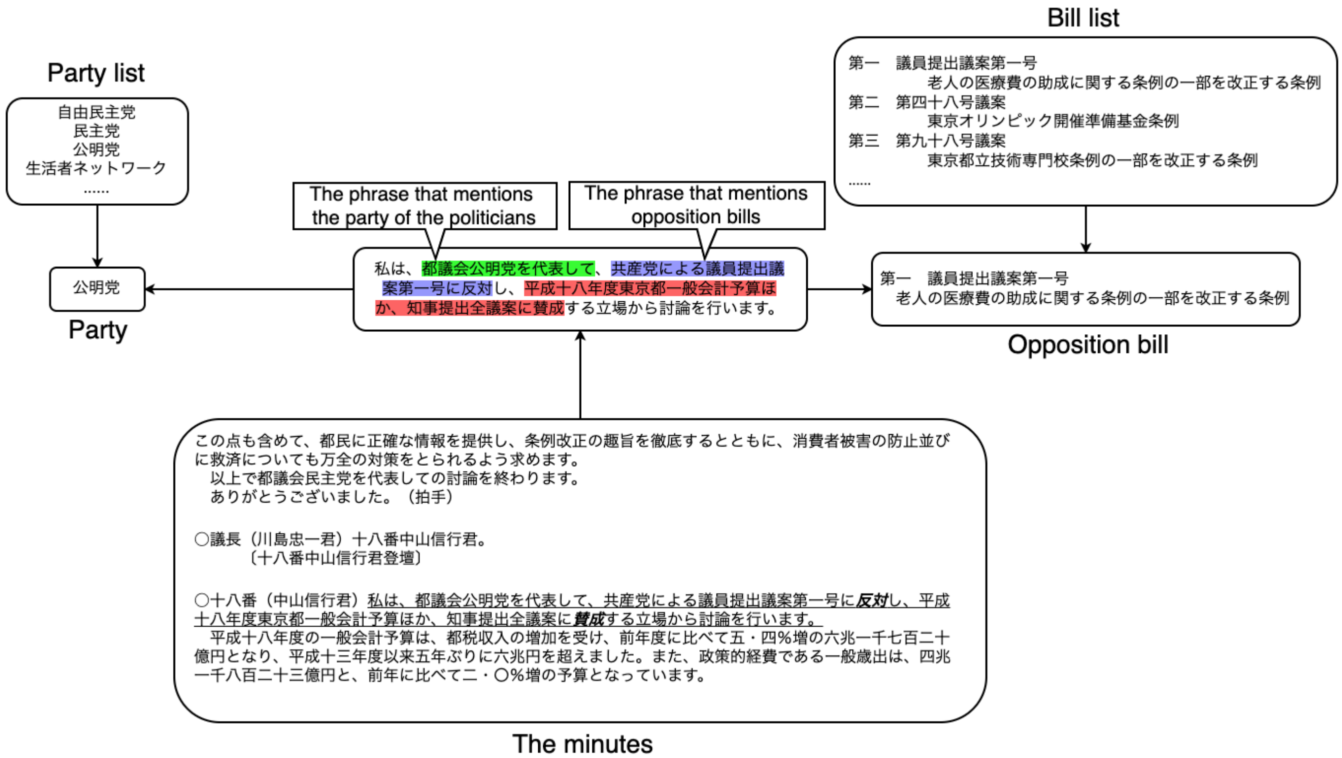

Figure 1: Stance classification from utterance

2.1.2 Committee utterances. Some parties state their stance against bills in the committee. We also detect stances from committee utterances by using the same method described in Section 2.1.1. In the committee, the politicians often mention detailed topics that have nothing to do with the stance of their party. For this reason, the rule-based method described in Section 2.1.1 sometimes results in incorrect stance detection. Therefore, we tightened the rules to prevent an incorrect extraction. In particular, we use the phrase "xxx 号議案" or "xxx 号の議案" instead of the phrase "xxx 号" to extract the bill number from the second and subsequent sentences of the utterance of the politician. In the committee, some politicians do not specify their parties in their utterances. In this case, we use a pre-prepared dictionary to look up the party where a politician belongs.

2.2 Classification stances from bill names

We can detect the stances of major parties using the method in Section 2.1. However, it is impossible to detect stances for some parties. We can categorize these cases into two.

(1) Some parties do not declare all the bills that they oppose because they oppose so many bills. Here is an example.

> "日本共産党都議団を代表して、第二百二十七号議案、都立老人医療センター条例を廃止する条例外二十一議案に反対し、議員提出議案、高齢者の医療費の助成に関する条例外二議案に賛成する立場から討論を行います。(On behalf of the Communist Party, I am going to debate from the standpoint of opposing 22 bills including bill No. 227, the Ordinance to Abolish the Municipal Medical Center for the Elderly, and supporting 3 bills submitted by city council members including the Ordinance to Subsidize Medical Expenses for the Elderly.)"

In this utterance, "日本共産党 (the Communist Party)" opposes 22 bills in total. However, it does not specify the other 21 bills (out of 22) that they oppose.

(2) Some small parties such as "市民の党 (the Citizen's Party)" and "自治市民 (the Autonomous Citizen Party)" rarely deliver an opinion in the meeting. In other words, there are few utterances from these small parties in the minutes of both the training and test sets. Therefore, it is impossible to detect a stance from the utterances of these parties.

Here, we do not discuss the former case because the rules in Section 2.3 can recognize stances. In this subsection, we describe a method for predicting a stance for the latter case. Each party has a tendency to oppose certain bills or topics. In other words, we can guess the stance of a party with no utterance only from a bill's name. Figures 2 and 3 provide an overview of the method. Here, we distinguish two types of bills: budget bills and normal bills. First, we explain the method for predicting a stance for a normal bill.

There are many instances of normal bills. Most of them only appear once in the training data. To extract the features from a bill name, we tokenize a bill name into n-grams and acquire the

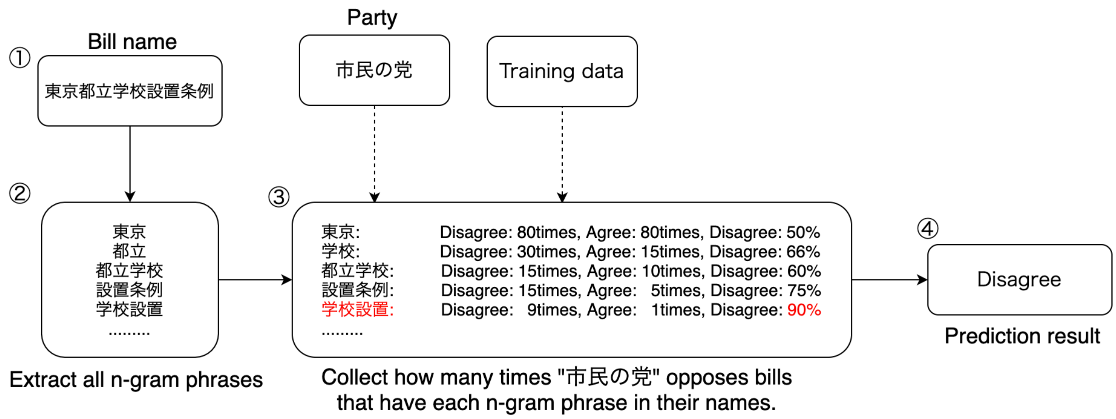

Figure 2: Stance classification from normal bill names

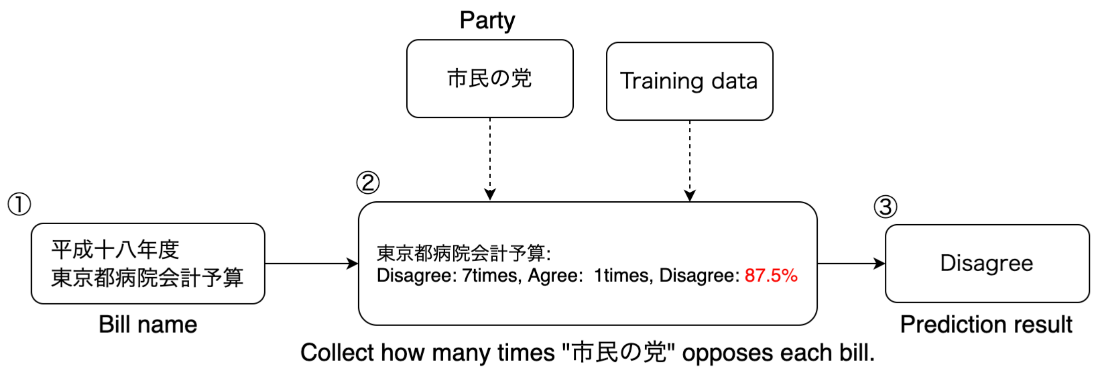

Figure 3: Stance classification from budget bill names

tendency of the stance of each party against the n-grams. We split a bill name into words by using MeCab [4]. For each n-gram in a bill name, we count the number of times each party opposes and agrees with the bills whose names include the n-gram. On the basis of the count and proportion of the opposition to and the agreement with an n-gram, we predict the stance of a party on a bill. Specifically, if a bill name includes an n-gram where the party agreed with the bill at least six times and never opposed it, we predict the stance of the party on the bill as an agreement. Similarly, if a bill name includes an n-gram where a party opposes the bill at least three times and opposes it 90% or more of the total bills, we predict the stance as an opposition[1].

For example, we tokenize the "東京都立学校設置条例 (Tokyo Metropolitan School Establishment Ordinance)" bill into "東京/都立/学校/設置/条例" and obtain n-grams, e.g., "東京," "都立," "都立学校," "設置条例," "学校設置," "東京都立学校," and "東京都立学校設置条例." If a party opposes bills 18 times and agrees with them 2 times, where the bills include "学校設置" in the training set, we predict the stance of this party on the bills including the n-gram as an opposition. If a bill name satisfies both the conditions for agreement and opposition, we predict the stance as an agreement.

We use a similar idea for budget bills. However, we do not apply tokenization for the bill names. This is because budget bills are discussed with the same name every year. Thus, we normalize a bill name by dropping the year from the name of a budget bill. We also treat the main budget bill and its supplemental budget bill as the same budget bill. For example, we normalize the bill "平成二十八年度東京都病院会計予算 (the fiscal year Heisei 28 Tokyo Hospital Account Budget)" into "東京都病院会計予算" and the bill "平成十四年度東京都一般会計補正予算（第一号）(the fiscal year Heisei 14 Tokyo General Account Supplementary Budget, first version)" into "東京都一般会計予算"[2]. We predict the stance of a party on a bill as an agreement when the party agreed to the normalized bill name at least twice and never opposed it. Similarly, we predict the stance of a party on a bill as an opposition when the party opposed the normalized bill name at least twice and opposed it 80% or more of the total bills. If a bill satisfies the conditions for both agreement and opposition, we predict the stance as an agreement.

[1]We tuned these parameters on the cross-validation in the training set.

[2]We can realize this process by using a simple rule-based algorithm.

As we explained earlier, we used a rule-based algorithm to predict stances from bill names. We also explored the use of ML-based methods to predict stances from bill names. However, the accuracy of the ML-based method was lower than that of the rule-based method. This is probably because of the insufficient supervision data for the ML-based method. Although the rule-based approach was sufficient to predict stances, we consider the use of the ML-based method in a future work.

2.3 Other clues

In addition to the method presented so far, we use other clues to improve the accuracy of stance detection. In this section, we briefly describe additional clues for stance detection.

2.3.1 Bills to be voted on at the same time. Multiple bills are usually voted on at a time, and party stances against those bills are always the same. Here is an example of the utterance of the chairman.

> "次に、日程第十三から第二十三まで、第六十号議案、学校職員の定数に関する条例の一部を改正する条例外議案十件を一括して採決いたします。(Next, we will collectively vote on Schedule No. 13 to No. 23, 11 bills including bill No. 60, the Ordinance to Amend Some Ordinances Relating to the Number of School Employees.)"

In this example, the stances of a party against the Schedule No. 13 to No. 23 bills are always the same. Therefore, once we predict the party's stance on one of these bills, we can predict its stance on the other bills.

2.3.2 Using the voting result to predict the stance. We can extract a voting result from the minutes. This result is useful for detecting stances. Here is an example of a voting result declared in an utterance.

> "次に、日程第十から第十二まで、第十五号議案、平成十八年度東京都用地会計予算外議案二件を一括して採決いたします。本案に関する委員会の報告は、いずれも可決であります。（中略）よって、本案は、いずれも委員会の報告のとおり決定いたしました。(Next, we will collectively vote on Schedule No. 10 to No. 12, three bills including bill No. 15, the fiscal year 2006 Tokyo Metropolitan Government's Land Use Account Budget. The committee reports that all bills were approved. (Omitted) Therefore, these bills have been decided as reported by the committee.)"

This utterance explains that the Schedule No. 10 to No. 12 bills are passed by a vote. The stances of most of the parties match the results of the vote. In other words, when a bill is passed, the governing parties usually agree with the bill. In contrast, when a bill is rejected, governing parties oppose the bill. We identify the (opposing) parties whose stances do not match the voting results from the training set. In this way, we associate the stances of the other parties with the voting results.

2.3.3 The condition where all parties agree with the bills. When the chair takes a vote, he/she speaks in either of these two patterns:

> (1) "本案は、起立により採決いたします。本案は、... 決定することに賛成の諸君の起立を求めます。(This vote will be taken on a standing vote. If you agree with the decision of ..., please stand up.)"
>
> (2) "お諮りいたします。本案は、... 決定することにご異議ありませんか。(Let me confirm. Are there any people who oppose the decision of ....)"

In the latter case, all parties agree with these bills. We extract this phrase and predict the stances of all parties on the corresponding bills as an agreement.

2.3.4 Minor opinion report. Some parties sometimes submit "少数意見報告書 (Minor opinion report)" in the plenary session when these parties oppose some bills when these bills were passed in the committee vote. When they submit this report, they always oppose these bills. We can identify the party names from the names of the politicians who submit the report.

2.3.5 Joint submission. Some parties submit bills jointly and state it in their utterance. Here is an example utterance for "共産党."

> "本定例会には、都議会維新の党、都議会生活者ネットワーク、かがやけ Tokyo、東京みんなの改革、そして日本共産党都議団の五会派の共同により、都議会議員の議員報酬、費用弁償及び期末手当条例一部改正条例案が上程されています。(In this plenary session, the Ordinance to Partially Amend the Ordinance on Remuneration, Expense Reimbursement, and End-of-term Allowances for Tokyo Metropolitan Assembly Members has been submitted, jointly submitted by five parties: the Restoration Party, the Seikatsusha Network Party, the Shine Tokyo Party, the Everyone Reform Party, and the Communist Party.)"

The stance of the parties who submit a bill is always in agreement with the bill. Therefore, we extract these parties who jointly submit the bill and regard their stance as an agreement.

2.3.6 Using the stances of other parties. After predicting the stance of a party, we can sometimes propagate the result to other parties. We describe two patterns.

(1) The stances of some small parties are likely to be the same as those of other major parties. When a major party that rarely opposes bills opposes a bill, it means that the bill is important and, thus, other small parties are likely to oppose it as well. This is true for approval. For example, when "民主党" opposes a bill, "市民の党" opposes the bill with a high probability. In particular, we find the following rule and use it:

(i) We predict the stance of "市民の党" and "自治市民" against the bill submitted by the governor as an opposition when one or more parties other than "共産党" and unaffiliated parties oppose the bill.

(ii) We predict the stance of unaffiliated parties that oppose the bill submitted by the governor at least once as an opposition when two or more parties other than "共産党" and unaffiliated parties oppose the bill.

(iii) We predict the stance of "ネット (the Network Party)," "みんな改革 (the Everyone Reform Party)," and "東京維新 (the Tokyo

Table 1: Statistics of the dataset

Dataset	Bill count	Stance count	Agreement count	Opposition count
Training	2,622	23,321	21,109 (90.5%)	2,212 (9.5%)
Test	479	4,541	4,265 (93.9%)	276 (6.1%)

Table 2: The result of the experiment

Other teams methods	Accuracy
Forst	93.88
akbl	94.98
knlab	95.31
Ibrk	96.50
Our methods	Accuracy
All Agreements	93.83
Utterances	96.43
Utterances + Committee utterances	97.39
Utterances + Committee utterances + Bill names	99.14
Utterances + Committee utterances + Bill names + Other clues	**99.75**

Table 3: The human evaluation result

Teams	Accuracy
Ibrk	-
knlab	83.4
Forst	85.2
akbl	89.2
Ours	**98.2**

Restoration Party)" on the bill submitted by the city council members as an agreement when two or more parties other than "共産党" and unaffiliated parties agree with the bill.

(2) The stances of the governing parties are always the same because these parties have made promises in advance. In fact, "自民党" and "公明党" were the governing parties from June 25, 2001, until July 1, 2017. "都民ファースト (the Tomin First Party)" and "公明党" have been governing parties since July 2, 2017. Once we can predict the stance of one of the governing parties, we can predict the stance of the other parties.

2.4 Overall process

We describe the overall process used to predict a stance. We apply the methods described earlier in the following order.

(1) The condition wherein all parties agree with the bills (Section 2.3.3)
(2) Minor opinion report (Section 2.3.4)
(3) Stance classification from plenary session utterances (Section 2.1.1)
(4) Joint submission (Section 2.3.5)
(5) Using the stances of other parties (Section 2.3.6)
(6) Stance classification from bill names (Section 2.2)
(7) Stance classification from committee utterances (Section 2.1.2)
(8) Using a voting result to predict stances (Section 2.3.2)

This order is given in descending order of accuracy of the stance detection. For example, the stance predicted by (1) is always correct. Once we predict a stance, we do not overwrite the stance in the later steps. Stances undetermined by any of the methods from (1) to (7) can be determined by method (8) because we can always extract a voting result. We also use the methods described in Section 2.3.1 every time we predict the stance using one of the methods from (1) to (7).

3 EXPERIMENTS

3.1 Dataset

We used the dataset distributed by the task organizers. The proceedings were taken from the official page of the Tokyo Metropolitan Assembly [3]. The stance of each party on each bill was obtained from "東京都議会だより" [4]. The dataset was split into the training and test sets, and the true stances in the test set were hidden during the task. We used cross-validation on the training set to tune the parameters. Table 1 summarizes the statistics of the dataset.

3.2 Evaluation metrics

In this task, our method was evaluated using two evaluation metrics. The first metric is the automatic evaluation metric. The predicted stance was evaluated by the percentage of agreement with

[3] https://www.gikai.metro.tokyo.jp/record.html
[4] https://www.gikai.metro.tokyo.jp/newsletter/

the gold stances in the test set. For the automatic evaluation metric, we predict the stance of each party as either an agreement or an opposition, as described in Section 2. The second metric is the human evaluation metric. In this metric, the stance extracted from only the utterance is evaluated. We predict the stance as either an agreement, an opposition, or not mentioned in the politician's utterance, and the other participants checked the part of the prediction manually. For the human evaluation metric, we made the output apart from the output for the automated evaluation metric. Specifically, we simply do not use any rules other than the rule of using the politicians' utterance described in Sections 2.1.1 and 2.1.2. The stance that cannot be predicted from utterances is predicted as "言及なし (not mentioned)." Refer to the overview paper [2] for details about the evaluation method.

3.3 Results

Table 2 reports the accuracy values of the proposed method as well as those of the other teams' methods. We included four variants of the proposed method: using utterances, committee utterances, bill names, and other clues. We used the clue described in Section 2.3.1 in all variants to reduce misclassification. Our method achieved 99.75% accuracy, which is 3.2 points higher than those of the other teams' methods. The table also indicates that all strategies in the proposed method contributed greatly to the accuracy. The accuracy achieved by our rule-based method was nearly 100%. This result indicates that this task was easy, probably because politicians express their opinions in a similar manner. This also suggests that there is room to make the task setting even more difficult. One possible direction is to prohibit the use of utterances that directly mention the stance of the politicians' party and, instead, to have the participants to acquire the potential stance of the party from its utterances.

Table 3 shows the result of the human evaluation metric. The proposed method also achieved the highest performance in this metric.

4 CONCLUSIONS

This paper presented our method for the stance classification task of NTCIR15 QA Lab-PoliInfo-2. There are two types of stances that we need to detect: stances that are explicitly stated and those that are not stated in the utterances. We designed a set of rules to recognize an explicit mention of a party's stance on a bill. When a party does not explicitly mention its stance, our method uses clues in the bill name to predict the party's stance. The proposed method achieved 99.75% accuracy in the automated evaluation metric and 98.2% accuracy in the human evaluation metric.

ACKNOWLEDGMENT

This work was supported by JSPS KAKENHI Grant Number 19H01118.

REFERENCES

[1] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2019. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*. 4171–4186.

[2] Yasutomo Kimura, Hideyuki Shibuki, Hokuto Ototake, Yuzu Uchida, Keiichi Takamaru, Madoka Ishioroshi, Teruko Mitamura, Masaharu Yoshioka, Tomoyoshi Akiba, Yasuhiro Ogawa, Minoru Sasaki, Kenichi Yokote, Tatsunori Mori, Kenji Araki, Satoshi Sekine, and Noriko Kando. 2020. Overview of the NTCIR-15 QA Lab-PoliInfo Task. *Proceedings of The 15th NTCIR Conference.*

[3] Elena Kochkina, Maria Liakata, and Isabelle Augenstein. 2017. Turing at SemEval-2017 Task 8: Sequential Approach to Rumour Stance Classification with Branch-LSTM. In *Proceedings of the 11th International Workshop on Semantic Evaluation (SemEval-2017)*. 475–480.

[4] Taku Kudo, Kaoru Yamamoto, and Yuji Matsumoto. 2004. Applying Conditional Random Fields to Japanese Morphological Analysis. In *Proceedings of the 2004 Conference on Empirical Methods in Natural Language Processing*. 230–237.

[5] Chi Sun, Luyao Huang, and Xipeng Qiu. 2019. Utilizing BERT for Aspect-Based Sentiment Analysis via Constructing Auxiliary Sentence. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*. 380–385.