Ibrk at the NTCIR-15 QA Lab-PoliInfo-2

Ryo Kato
Department of Computer and Information Sciences
Faculty of Engineering/Ibaraki University
4-12-1, Nakanarusawa, Hitachi, Ibaraki, 316-8511, Japan
17t4025h@vc.ibaraki.ac.jp

Minoru Sasaki
Department of Computer and Information Sciences
Faculty of Engineering/Ibaraki University
4-12-1, Nakanarusawa, Hitachi, Ibaraki, 316-8511, Japan
minoru.sasaki.01@vc.ibaraki.ac.jp

ABSTRACT

In this study, we develop a system that automatically identifies whether each party agrees or disagrees with each bill on the minutes of the Tokyo Metropolitan Assembly. For the interpellations given by the members of each party and their answers in the minutes of the meeting, we predict whether the members of each party are for or against each bill. It is not difficult to predict agrees or disagrees with a bill if members have given their opinions on it. However, it is difficult to predict whether the minority parties agree or disagree with the bill because the minority parties have little opportunity to give speech. In this paper, we propose a method for predicting agree or disagree with a bill based on the external knowledge and the past meeting proceedings when it is not possible to extract agree or disagree with a bill.

Team Name

• Ibrk

Subtasks

• Stance Classification Task (Japanese)

KEYWORDS

Stance classification, Minutes of the Tokyo Metropolitan Assembly, Minority Party

1 Introduction

In this paper, we develop a system that allows each party to decide whether they agree or disagree with an individual bill in a local assembly.

Currently, a number of local governments make available to the public the minutes of the meetings of their local assemblies, as their efforts to disclose information. In order to encourage research on local politics and its applications, a corpus of local assembly minutes is developed that researchers can use [1]. In this context, various studies have been conducted using these corpora, such as information extraction and visualization of assembly activities.

In this study, we develop a system that automatically identifies whether each party agrees or disagrees with each bill on the minutes of the Tokyo Metropolitan Assembly. For the interpellations given by the members of each party and their answers in the minutes of the meeting, we predict whether the members of each party are for or against each bill. It is not difficult to predict agrees or disagrees with a bill if members have given their opinions on it. However, it is difficult to predict whether the minority parties agree or disagree with the bill because the minority parties have little opportunity to give speech. In this paper, we propose a method for predicting agree or disagree with a bill based on the external knowledge and the past meeting proceedings when it is not possible to extract agree or disagree with a bill.

2 Stance Classification System

2.1 System Overview

We scan the minutes of each regular meeting to detect "agree" or "disagree" of the agenda from statements. If we are able to detect them, we take their approval or disapproval as the opinion of the speaker's faction. In the event that we do not obtain an approval or disapproval of the proposal being discussed for each faction, we assign an opinion according to the rule base we have developed. We will take the "agree" or "disagree" thus obtained as our response.

2.2 Text entry and Preprocessing

We use the meeting minutes of the plenary session. We collect the date of the last day of the meeting from the Tokyo Metropolitan Assembly's website (https://www.gikai.metro.tokyo.jp/record/proceedings/) where the last day of the meeting can be obtained. We scanned the proceedings by sequence and evaluated the proceedings by sequence. Speeches are described in the order of speakers and their content (Utterance), grouping them together.

If the value of the speaker is null and the meeting is on the last day, the members present and the agenda to be discussed are described in the Utterance. The number of the agenda items is obtained in order. If the value of Speaker is not null, we extract only the speaker's name and store them in the order of appearance.

We compare the stored speaker's name with the speaker's dictionary and determine the speaker's party affiliation based on the meeting date and name. By doing so, we extract the opinion of the group from the speaker.

2.3 How to extract and approve or reject bills

We divide Utterance into sentences. We use "MeCab" to morphologically analyze these sentences. We use the IPADIC in

MeCab.This yields a list of strings. We scan this list to see if there are "agree" or "disagree" word. If we find either of them, we divide the string into two based on the part of the phrase. We determine that if the previous sentence, divided into two parts, states "Bill No. ○○" and the proposition number, we determine that the proposition is the opinion of the found language. The same action is then repeated in the back sentence. When we evaluate the previous sentence, it may be described as "Outside ○○ bill". In this case, we convert ○○ into alphanumeric characters and determine that the number of propositions obtained from the top of the agenda item + 1, out of all the propositions obtained when the speaker was null in 2.2, is considered to be a one opinion. For those proposals that do not receive a favorable or unfavorable opinion from each party, the Communist Party, which is often the opposition party, is given an opposing opinion, and other parties are tentatively given an affirmative opinion.

2.4 Decisions for and against by the ruling and opposition factions

In 2.3, we assigned affirmative opinions to the proposals on which no opinion was obtained. We now assign the labels of the ruling and opposition parties to the minority caucus and assign opinions accordingly. The minority group has very little to say during the assembly and it is difficult to extract approval or disapproval from the text of the meeting. For this reason, we collect the tendency of each minority group to agree or disagree with the opinion of the ruling party or the opposition party in advance, and assign opinions for or against each minority group accordingly.

We estimated this tendency from Wikipedia's Tokyo Metropolitan Assembly page (https://ja.wikipedia.org/wiki/%E6%9D%B1%E4%BA%AC%E9%83%BD%E8%AD%B0%E4%BC%9A) and training data, and then we used the following data , a dictionary was created. We grant approval or disapproval if no approval or disapproval is obtained. For minority parties, we refer to the dictionary we created. A faction that is presumed to have "oppositional tendencies" is assigned the same opinion as the Communist Party, which is often the largest opposition party, as the opinion of that faction. This assigns an opinion to the minority party.

2.5 Estimation of Trends for and against the various factions

We consider, as in 2.4, the manner in which opinion is presumed for the minority group. At the first meeting of each year of the meeting minutes, there is a discussion of the proposed budget.

This budget proposal includes the "General Fund" and "Water Utility Account" and often the same budget proposal is discussed each year. We also estimate that all political parties often have a same opinion on this budget proposal every year. We gather from the Training data which budget proposals the various factions disagree with each year and create a dictionary. Using this dictionary, we determine whether the parties in question have ever held opposing views on the budget proposal, and if so, we assign opposing views to it as well.

3 Experiment

3.1 Date

The data used are the minutes of the Tokyo Metropolitan Assembly meetings (from the fourth regular meeting in 1999 to the first regular meeting in 2019) and the test data (4,551 questions) given in the stance classification task. Training data entry file (23,321 questions). This question file contains information about the meeting and the name of the agenda. We apply it to the information of each faction for or against the agenda item being discussed in the meeting, and get the percentage of correct answers.

3.2 results

We conduct the experiment in four ways:

(1) when no approval or disapproval of the governor's proposal is obtained, the Communist Party is assigned an opposing opinion and the other factions are assigned an opinion in favor of the proposal;

(2) when no approval or disapproval of the governor's proposal is obtained, all factions are assigned an opinion in favor of the proposal;

(3) In addition to (2), information on the tendencies of the ruling and opposition parties is assigned to the minority caucus;

(4) In addition to (2), information on the minority caucus' budget proposal is assigned.

3.2.1 Granting a dissenting opinion to the Communist Party in the governor's proposal

We used the method described in 2.2 to verify the accuracy. In the training data, out of 23,321 questions, the number of correct answers was 20,387, the number of wrong answers was 2,934, and the accuracy was 87.42%.In the Test data, there were 4,551 questions, of which 4,028 were correct, 513 were wrong, and the accuracy was 88.70%. Of the Communist Party opinions in the Test data, 268 were labeled with mention and 211 were labeled without mention. Among the unmentioned labeled opinions, 196 were correct and 15 were incorrect, for a correct response rate of 92.89%.

3.2.2 If no approval or disapproval is received, all are given an affirmative opinion

In 2.3, we assigned a dissenting opinion to the Communist Party on the gubernatorial proposal that did not receive approval or disapproval, but this time we assigned approval to all of the gubernatorial proposals that did not receive approval or disapproval to test the accuracy. In the Training data, out of 23,321 questions, the number of correct answers was 22,042, the number of wrong answers was 1,279, and the accuracy was 94.52%; in the Test data, out of 4,551 questions, the number of correct answers was 4,361, the number of wrong answers was 180, and the accuracy was 96.04%.

3.2.3 Granting information on the ruling and opposition parties in the minority party

The method described in 2.4 was added to the method in 3.2.2 to verify the accuracy. We have designated the various parties as other than 「自民党」,「日本共産党」,「公明党」,「民進党」,「民主党」,「都民ファースト」. In the training data, the number of correct answers was 22,115, the number of wrong answers was 1,206, and the accuracy was 94.83% out of 23,321 questions. In the Test data, out of 4,551 questions, the number of correct answers was 4,377, the number of wrong answers was 164, and the accuracy was 96.39%. Of the opinions on the Test data, 642 were labeled with mention and 1984 were labeled without mention. Among the unmentioned labeled opinions, there were 1854 correct answers and 130 incorrect answers, for a correct response rate of 93.45%.

3.2.4 Granting information on the minority group's budget proposal

The method described in 2.5 was added to the method in 3.2.2 to verify the accuracy. In the training data, of the 23,321 questions, 22,042 were correct, 1,279 were wrong, and the accuracy was 94.52%. In the test data, of the 4,551 questions, 4,361 were correct, 180 were wrong, and the accuracy was 96.04%.

	3.2.2	3.2.3
No mention	211	642
Mentioned	268	1984
sum	479	2626

Table 1:Number of mentions and none in 3.2.2 or 3.2.3

	3.2.2	3.2.3
miss	15	130
match	196	1854
accuracy	0.92891	0.934476

Table 2:Percentage of correct answers with no mention in 3.2.2 or 3.2.3

3.3 Comparative Rating

The results of the experiment are shown in Table 3.

We predicted that under method (1), the Communist Party would often oppose the Governor's proposal because of the opposition's position.

Comparing the results of (1) with (2), we found a 7.34% improvement in accuracy. This is a result of the Communist Party also often granting an opinion in favor of the Governor's proposal.

In (3), we predicted the tendency of the minority party to not be able to speak and assigned an opinion. As a result, we found a 0.35% improvement over (2) and obtained the highest performance.

The result of (4) was the same as the result of (2).

The percentage of correct answers was 96.39% for the combined functions of (3) and (4).The results showed that the performance of (2) and (3) overlapped, or the usefulness of (4) was low.

	(1)	(2)	(3)	(4)
miss	513	180	164	180
match	4028	4361	4371	4361
accuracy	88.70%	96.04%	96.39%	96.04%

Table 3: Comparison of performance for each rule base

4 Conclusion

In this paper, we proposed a method for estimating the approval or disapproval of a bill from statements in the Tokyo Metropolitan Assembly minutes. We tested three estimation techniques and compared their performance. One examined the method of assigning opinions when opinions were not obtained. The second was the estimation of opinions using the dictionary of the ruling and opposition parties for the minority group. Finally, we created a dictionary on the minority party's budget proposal and estimated opinions based on it. The accuracy of each of these methods was examined to determine which method would be more accurate. As a result, the model with the highest accuracy was the model of using the dictionary of the ruling and opposition parties for the minority caucus, which provided a high performance of 96.39% of correct answers.

REFERENCES

[1] Yasutomo Kimura, Keiichi Takamaru, Takuma Tanaka, Akio Kobayashi, Hiroki Sakaji, Yuzu Uchida, Hokuto Ototake and Shigeru Masuyama, 2016. Creating Japanese Political Corpus from Local Assembly Minutes of 47 Prefectures. *Proceedings of the 12th Workshop on Asian Language Resources (ALR12)*, 78-85.

[2] Sten Andler. 1979. Predicate path expressions. In *Proceedings of the 6th. ACM SIGACT-SIGPLAN Symposium on Principles of Programming Languages (POPL '79)*. ACM Press, New York, NY, 226-236. DOI:https://doi.org/10.1145/567752.567774

[3] Ian Editor (Ed.). 2007. *The title of book one* (1st. ed.). The name of the series one, Vol. 9. University of Chicago Press, Chicago. DOI:https://doi.org/10.1007/3-540-09237-4.

[4] David Kosiur. 2001. *Understanding Policy-Based Networking* (2nd. ed.). Wiley, New York, NY..

[5] 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.

[6] 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.