fuys at the NTCIR-16 QA Lab-PoliInfo-3 Budget Argument Mining Subtask

fuys team

Daigo Nishihara

Hokuto Ototake

Kenji Yoshimura

Fukuoka University
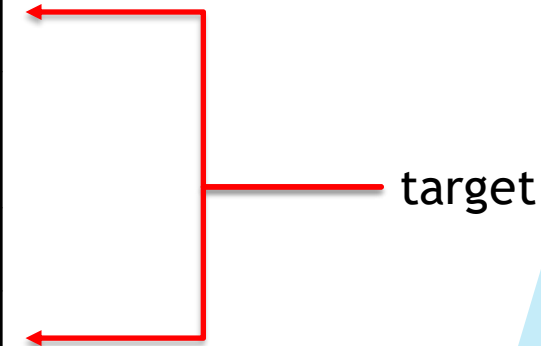
PROPOSED METHOD（ArgumentClass）

▶ Model and infer using BERT.
（1248 monetary expressions covered）

▶ Utterances is divided by "、(comma)", "。(period)" and "space".
（Less than 15 characters are combined with the previous sentence.）

▶ The target sentences are those in which monetary expressions.

▶ The label of our model is ArgumentClass of the monetary expression.

PROPOSED METHOD（ArgumentClass)

Examples of delimiters and target sentences

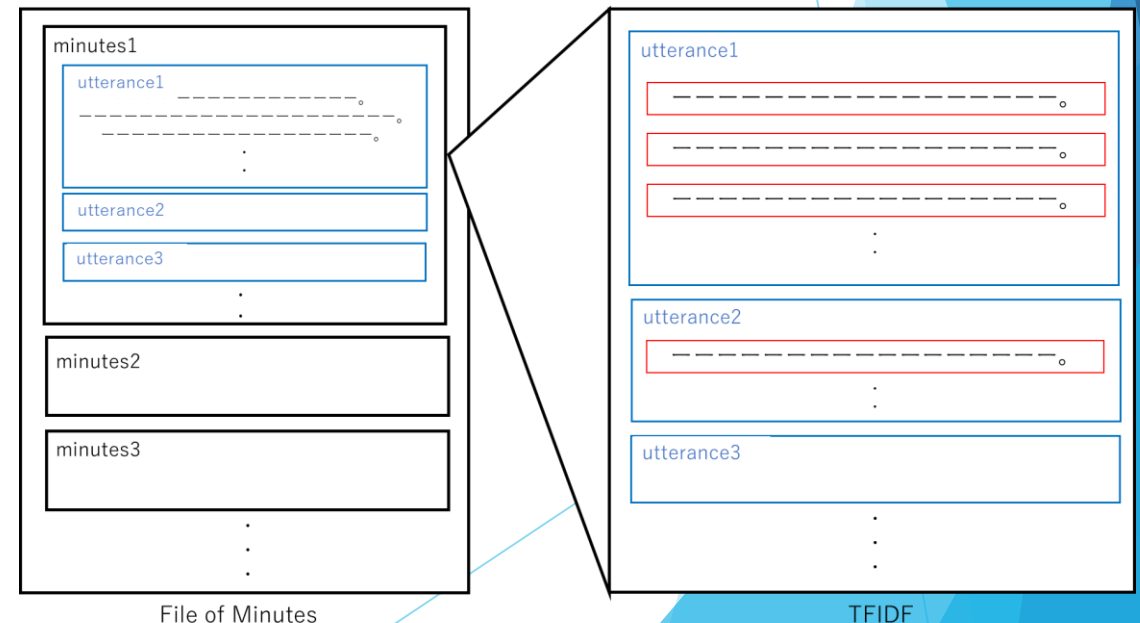| After division	国民健康保険加入者は退職者など高齢者が増え、３割を占めており、また、非課税所得者が約５割、
	所得割保険料を負担する世帯の約７割が所得額200万円以下となっていることから、
	所得割負担者の負担がより重くなる構造になっています。
	消費税が引き上げられ、
	生活は一段と苦しくなっていく中で、年間2,000円の負担増は重く、
	保険料値上げには賛同できません。
	今回の財政調整基金設置の背景にある国の法定外繰入れ削減強要による被保険者への負担増は認めることはできません。

target

PROPOSED METHOD （ArgumentClass)

► Labels are biased depending on the speaker.

→ Some meetings don't get "Claim" on the non-legislator's label.

► Additional dimensions for determining whether a person is a legislator or not. (giin_flag)

Speaker's Position	giin_flag
legislator	1
No legislator (e.g., Mayor, cabinet minister ,etc.)	0

PROPOSED METHOD（RelatedID）

▶ Assigning using key word extraction with TFIDF.

▶ key words are nouns, not numerals, with at least 3 letters.

▶ Sentence units that further divide the utterances in the assembly minutes.

▶ Utterances is divided by "。(period)" and "space".

(Less than 200 characters are merged into the sentence behind.)
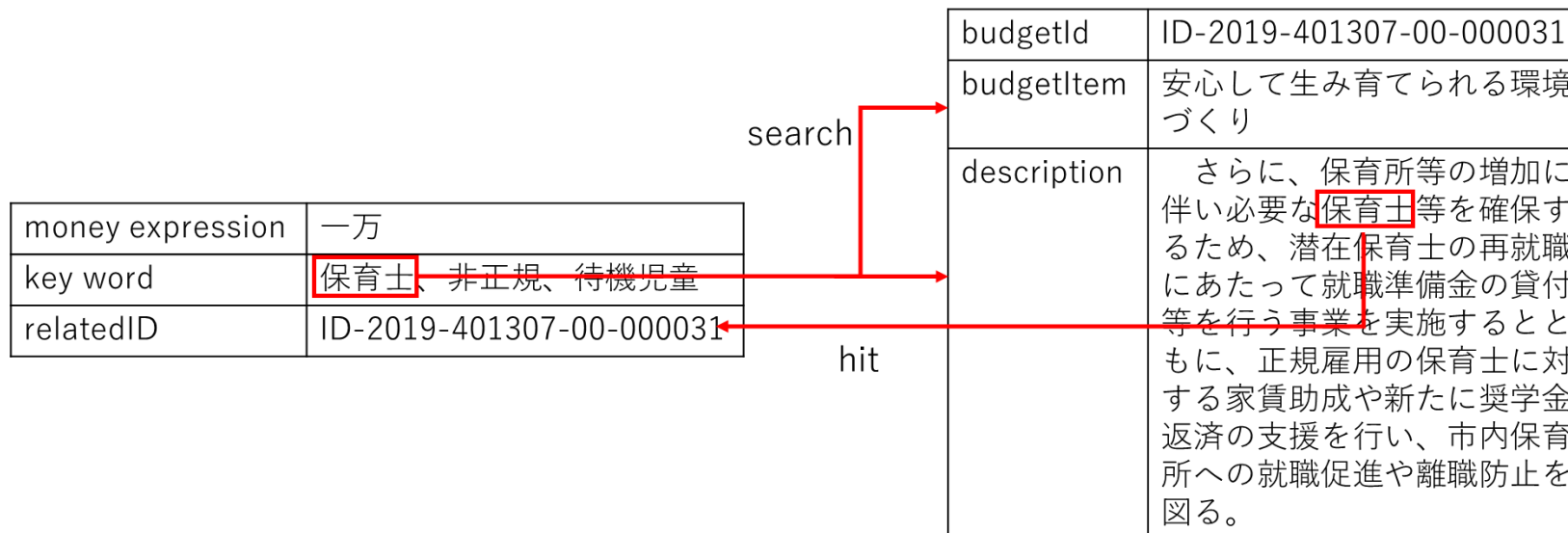

File of Minutes   TFIDF

PROPOSED METHOD（RelatedID）

▶ The words with the highest TFIDF value in the sentence are looked at in order.

▶ The budget items to look for are looked at from the youngest number in the budgetId.

money expression	一万
key word	保育士、非正規、待機児童
relatedID	ID-2019-401307-00-000031

search

hit

budgetId	ID-2019-401307-00-000031
budgetItem	安心して生み育てられる環境づくり
description	さらに、保育所等の増加に伴い必要な保育士等を確保するため、潜在保育士の再就職にあたって就職準備金の貸付等を行う事業を実施するとともに、正規雇用の保育士に対する家賃助成や新たに奨学金返済の支援を行い、市内保育所への就職促進や離職防止を図る。

RESULTS（test data）

	No "giin-flag"			"giin-flag"		
	local	diet	all	local	diet	all
score	23.91%	0%	23.40%	23.91%	0%	23.40%
ArgumentClass	<mark>59.56%</mark>	38.46%	<mark>56.92%</mark>	55.38%	<mark>41.54%</mark>	53.65%
RelatedID	34.78%	0%	34.04%	34.78%	0%	34.04%

CONSIDERATION（ArgumentClass）

▶ There was no difference in scores with and without the "giin-flag".

▶ The percentage of correct answers differs by only less than 5% points.

→ No advantage is seen in the correctness rate with and without the "giin_flag".

	No "giin-flag"		"giin-flag"	
	local	diet	local	diet
score	23.91%	0%	23.91%	0%
ArgumentClass	59.56%	38.46%	55.38%	41.54%

CONSIDERATION（RelatedID）

▶ Accuracy is poor.

▶ Causes of failure to improve scores

▶ There are problems with our proposed method.

score	23.40%
RelatedID	34.04%