

*nagoy* Team's  
Summarization System  
at the NTCIR-14  
QA Lab-PoliInfo

Yasuhiro Ogawa,  
Michiaki Satou,  
Takahiro Komamizu,  
Katsuhiko Toyama  
(Nagoya University)

# Motivation

---

---

- Summarization for Japanese statutes
  - using **Random Forest Classifier**
  - based on **sentence extraction**
  - good result
  
- Our method is also useful for assembly member speeches?

# Problems

---

---

1. Correct summary is not made by sentence extraction
  - ← random forest is based on sentence extraction
2. Imbalanced data
  - ← different imbalance form statute summary
3. Length limit
  - ← random forest is a binary classifier

# Problems

Togikai dayori

1. Correct summary is **not** made by **sentence extraction** from the source document

← random forest is based on sentence extraction

Assembly members speeches

2. Imbalanced data

← different imbalance form statute summary

How to make correct data (training data)?

← random Forest is a binary classifier

# How to Make Training Data

---

---

- **Positive sentence**: used in the summary
  - speech sentence that contains most words in the summary
  - if the speech sentence is shorter than the limit, we add the second-most sentence
- **Negative sentence**: not used in the summary
  - otherwise

# Training Data

---

---

	Formal-run
documents	596
sentences	9979
positive sentences	825 (8.3%)

Imbalanced



different

statute summarization: 70%

# Problems

---

---

1. Correct summary is not made by sentence extraction

making training data choosing similar one

2. Imbalanced data

progressive ensemble random forest <sup>ary</sup>

3. Length limit

← Random Forest is a binary classifier

# Undersampling

Which classifier should we use?

Document ID	111	106	23	19	92		
# of sente	All classifiers step by step						
	N P	1	1	9	3	7	5
	N N P	2	0	5	2	3	3
	N N N P	3	0	2	1	3	1
	N N N N P	4	0	0	0	1	1
	N N N N N P	5	0	0	0	0	1



*progressive ensemble random forest*



# Problems

---

---

1. Correct summary is not made by sentence extraction

making training data choosing similar one

2. Imbalanced data

progressive ensemble random forest

3. Length limit

sentence reduction binary classifier

# Sentence Reduction

---

---

活気のある社会を構築すべきと考えますが知事の所見を伺います。

1. Preprocessing:  
Replacing the end of sentences

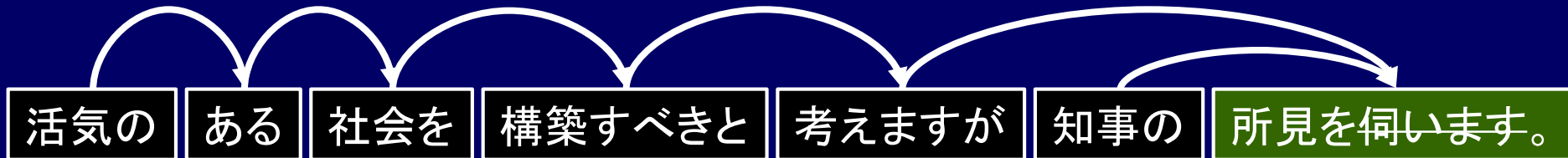
# Preprocessing: String Replacement

target string	replaced string
でございます。	です。
伺います。	。
ております。	ている。
でおります。	でいる。
てまいります。	ていく。
でまいります。	でいく。
であります。	です。
いたします。	する。
と思います。	。
と思っている。	。

# Sentence Reduction

---

---



1. Preprocessing:  
Replacing the end of sentences
2. Analyzing dependency by CaboCha
3. Choose the last segment (= *bunsetsu*)

# Sentence Reduction

活気 <small>の</small>	ある	社会を	構築すべきと	考えますが	知事 <small>の</small>	所見を伺います。
100	200	300	400	500	500	
40	0	20	0	0	40	
0	0	13	108	0	61	
140	200	333	508	500	601	

## 4. Scoring each segment

- dependency depth
- case information
- frequency in all summaries

# Sentence Reduction



活気の	ある	社会を	構築すべきと	考えますが	知事の	所見を伺います。
100	200	300	400	500	500	
40	0	20	0	0	40	
0	0	13	108	0	61	
140	200	333	508	500	601	

5. Adding the segment with the highest score
  - adding the other segments on the path
6. Stopping when exceeds the limit

# EXPERIMENTS

# Random Forest Features

---

---

- Sentence position
- Sentence length
- Presence of a word
  - noun
  - occurs more than once in the summaries
  - not within the top 20 occurrences in the source



# Experimental Setting

---

---

- Training data
  - 596 documents 9,979 sentences
- Test data
  - 146 documents
- Random Forest by scikit-learn
  - 992 words as features

# Quality Question Scores

	all-topic				single-topic				multi-topic			
	content		formed	total	content		formed	total	content		formed	total
	X=0	X=2			X=0	X=2			X=0	X=2		
KitAi-01	0.856	1.134	1.732	0.912	0.953	1.170	1.660	0.995	0.745	1.092	1.815	0.815
KitAi-02	0.788	1.035	1.308	0.667	0.849	1.028	1.340	0.722	0.717	1.043	1.272	0.603
TTECH-01	0.290	0.644	1.783	0.402	0.274	0.575	1.755	0.401	0.310	0.723	1.815	0.402
nagoy-01	0.886	1.104	1.619	0.899	0.953	1.179	1.642	1.028	0.810	1.016	1.592	0.750
akbl-01	0.722	1.005	1.833	0.826	0.708	1.009	1.844	0.849	0.739	1.000	1.821	0.799
akbl-02*	0.707	1.000	1.837	0.793	—	—	—	—	0.707	1.000	1.837	0.793
KSU-01	0.043	0.043	1.955	0.048	0.052	0.052	1.934	0.057	0.033	0.033	1.978	0.038
KSU-02	0.076	0.121	1.745	0.071	0.080	0.156	1.722	0.104	0.071	0.082	1.772	0.033
KSU-03	0.091	0.157	1.715	0.104	0.104	0.179	1.731	0.156	0.076	0.130	1.696	0.043
KSU-04	0.111	0.167	1.419	0.093	0.118	0.193	1.420	0.132	0.103	0.136	1.418	0.049
KSU-05	0.048	0.078	1.692	0.048	0.057	0.085	1.726	0.057	0.038	0.071	1.652	0.038
KSU-06	0.078	0.169	1.535	0.091	0.085	0.151	1.542	0.094	0.071	0.190	1.527	0.087
LisLb-01	0.720	0.942	1.237	0.591	0.722	0.920	1.349	0.684	0.717	0.967	1.109	0.484
TO-01	0.504	0.846	1.763	0.551	0.464	0.794	1.778	0.521	0.550	0.905	1.746	0.586
average	0.423	0.603	1.655	0.435	0.387	0.535	1.532	0.414	0.406	0.599	1.646	0.394

formed score is low: bad sentence reduction

content score is high: good sentence extraction

# ROUGE Scores (Surface Form)

	recall							F-measure						
	N1	N2	N3	N4	L	SU4	W1.2	N1	N2	N3	N4	L	SU4	W1.2
KitAi-01	0.440	0.185	0.121	0.085	0.375	0.217	0.179	0.357	0.147	0.096	0.067	0.299	0.168	0.188
KitAi-02	0.390	0.174	0.113	0.078	0.320	0.200	0.154	0.343	0.154	0.101	0.069	0.281	0.173	0.176
TTECH-01	0.278	0.060	0.035	0.020	0.216	0.092	0.096	0.240	0.055	0.031	0.018	0.187	0.079	0.111
nagoy-01	0.459	0.200	0.131	0.089	0.394	0.229	0.186	0.361	0.151	0.097	0.064	0.305	0.169	0.192
akbl-01	0.400	0.173	0.113	0.076	0.345	0.189	0.157	0.361	0.156	0.102	0.068	0.310	0.167	0.185
akbl-02	0.326	0.124	0.080	0.057	0.269	0.147	0.112	0.320	0.119	0.077	0.055	0.262	0.141	0.144
KSU-01	0.158	0.028	0.009	0.002	0.147	0.043	0.071	0.210	0.039	0.013	0.004	0.196	0.059	0.107
KSU-02	0.185	0.043	0.021	0.014	0.167	0.063	0.080	0.230	0.056	0.027	0.017	0.209	0.080	0.116
KSU-03	0.172	0.036	0.008	0.002	0.157	0.050	0.075	0.211	0.043	0.011	0.003	0.192	0.062	0.106
KSU-04	0.171	0.044	0.013	0.002	0.153	0.055	0.072	0.219	0.056	0.017	0.003	0.195	0.072	0.106
KSU-05	0.227	0.029	0.010	0.002	0.195	0.064	0.089	0.231	0.029	0.010	0.003	0.196	0.065	0.110
KSU-06	0.221	0.038	0.013	0.004	0.187	0.065	0.086	0.230	0.038	0.012	0.004	0.192	0.067	0.108
LisLb-01	0.251	0.120	0.079	0.058	0.211	0.132	0.103	0.226	0.107	0.071	0.051	0.188	0.115	0.118
TO-01	0.267	0.093	0.061	0.045	0.230	0.117	0.105	0.272	0.086	0.052	0.036	0.233	0.110	0.133

good sentence extraction

random forest is good

# Successful Example

次に、新しい公共について伺います。¥¥n新しい公共という考え方は、私たちが国家戦略の柱として、地域主権改革とともに、これからのあるべき社会像として掲げたものです。¥¥n日本では、古くから連、結、講、座、あるいは若者組などの住民組織や市井の寺子屋、隠居という名のボランティア的な活動などが活力ある市民社会を担っていました。¥¥n新しい公共の考え方は、以前あったこのような社会を現在にふさわしい形で再構築することを目指すものです。¥¥n東日本大震災の被災地では、数々のボランティア活動が行われています。¥¥n強制ではなくみずからの意思で支援活動をされていた多くの方々の姿は感動的であり、改めて人々のつながりと助け合いの大切さを感じさせられました。¥¥n石原都知事は、都の防災対応指針において、自助、共助の徹底について述べられています。¥¥n行政依存ではなく、一人一人自立した個が、地域、社会を主体的に働きかけていく協働は、災害時には不可欠なものです。¥¥nそこで伺います。¥¥n東京都においては、このような新しい**公共型社会の実現を目指し、支え合いと活気のある社会を構築していくべきと考えますが、知事の所見を伺います。**¥¥n

GS: 支え合いと活気のある社会を構築すべき。知事の所見を。

公共型社会の実現を目指し、支え合いと活気のある社会を構築していくべきと考えますが、知事の所見を。

# Unsuccessful Example

我が国の経済にあっては、欧州の債務危機や歴史的な円高などが、回復の兆しが見えた景気に冷や水を浴びせています。¥¥n企業収益の動向は不透明さを増しており、今後の都税収入への影響は避けられません。¥¥nこうした中、都には、少子高齢化や中小企業対策など、山積する課題に対して効果的な手だてを講じ、現下の閉塞感を打ち破り、東京に活力を呼び戻していくことが求められています。¥¥nとりわけ、震災への対応は喫緊の課題です。¥¥n我が党が立ち上げた東日本大震災復旧・復興対策推進本部で議論を重ね、先月、防災力強化に向けての提言を行いました。¥¥n提言内容も含め、高度防災都市の実現に向けた取り組みを加速する上では、法人事業税の暫定措置の撤廃は不可欠であり、約束どおり撤廃するよう国に強く求めるものであります。¥¥nこの間、国が公共事業を見識ある考えもなく削減し続けたのとは対照的に、都は七年連続で投資的経費を伸ばしてきました。¥¥n都税収の回復が当面期待できない今だからこそ、中小企業の受注機会をふやすなど、景気を刺激し、防災力強化にも資する投資的経費に財源を振り向けることが重要であります。¥¥nこれまで以上にめり張りをつけ、都民に安心と希望をもたらす予算とするべく、新年度予算編成作業を進めるべきと考えますが、所見を伺います。

GS:メリハリをつけて、都民に安心と希望をもたらす予算にするべき。  
所見は。

つけ、都民に安心と希望をもたらす予算とするべく、新年度予算編成作業を進めるべきと考えますが、所見を。

# DISCUSSION

# Random Forest with Probability

---

---

- Proposed method (**progressive ensemble RF**) uses multiple classifiers step by step
- Scikit-learn's random forest
  - outputs results with their **probability**
  - is able to choose the sentence with the **highest probability**

Which is better?

# Comparison with using Probability

using the same model but outputs the highest-probability sentence

	Precision				Recall				F-measure			
	closed	open			closed	open			closed	open		
		all	single	multi		all	single	multi		all	single	multi
1	0.860	0.465	0.482	0.450	0.785	0.437	0.432	0.441	0.821	0.450	0.456	0.446
2	0.967	0.471	0.464	0.477	0.893	0.406	0.411	0.402	0.929	0.436	0.436	0.436
3	0.973	0.520	0.560	0.483	0.875	0.452	0.495	0.412	0.921	0.484	0.525	0.444
4	0.983	0.526	0.588	0.466	0.886	0.462	0.526	0.402	0.932	0.492	0.556	0.432
5	0.987	0.511	0.553	0.473	0.886	0.457	0.495	0.422	0.933	0.483	0.522	0.446
all	1.000	0.523	0.571	0.477	0.896	0.457	0.505	0.412	0.945	0.488	0.536	0.442
proposed	0.963	0.446	0.481	0.417	0.967	0.523	0.526	0.520	0.965	0.481	0.503	0.463

- Our method scored the highest in recall
- Our method does not need to tune the balance of positive and negative data



# Conclusion

---

---

- Summarization task using **progressive ensemble random forest**
  - good results in ROUGE Scores
- Future task
  - improve the sentence reduction module
  - investigate the relationship between the proposed method and the RF with probability