

# AMI Team at the NTCIR-16 Real-MedNLP Task

Satoshi Hiai, Shoji Nagayama, and Atsushi Kojima  
Advanced Media, Inc.



Advanced Media, Inc.

- Medical Named Entity Recognition

**A**左上葉に**F**径18mm **F**大の**D(+)**SSN を認めます。**D(?)**AAH や**AIS** の可能性があります。**A**右下葉にも**D(+)**GGN を**F**散見 します。**D(?)**炎症性変化 かもしれませんが、フォローにて**D(+)**変化 をご確認ください。**A**左下葉に**D(+)**線状索状影 を認め**D(?)**陳旧性炎症性変化 が疑われます。**A**縦隔や肺門に**F**有意な**D(-)**リンパ節腫大 は指摘できません。**D(-)**胸水 はありません。

- Subtask1
  - The organizer provides approximately 200 documents
- Subtask2
  - Participants construct NER methods based on the guideline
  - The guideline contains a handful of sample sentences

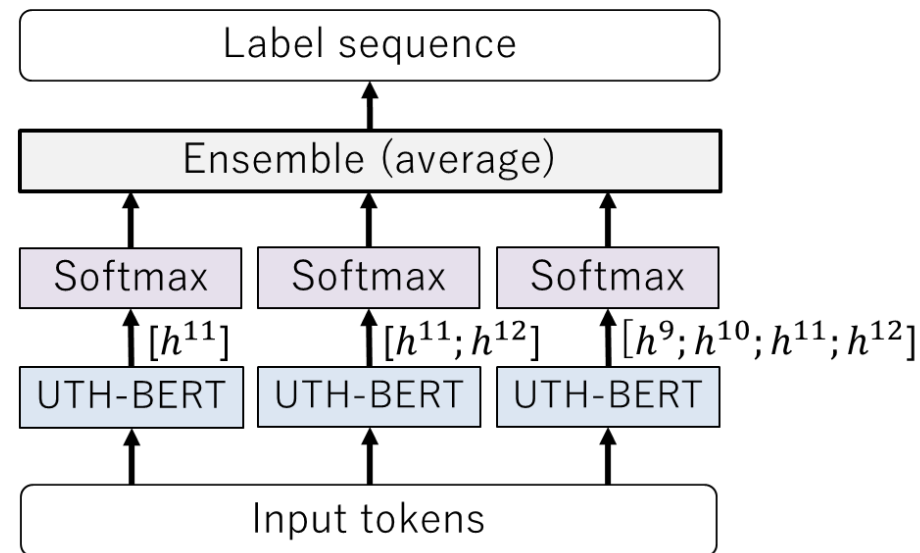
Need for methods for few resources



- Subtask1: **BERT-based approach**
  - Pretrained language models have achieved excellent results for low resource situations
  - We tried two methods with BERT pretrained on biomedical documents (UTH-BERT)
- Subtask2: **Rule-based approach**
  - Used rule-based approach due to Insufficient data for machine learning methods
  - Used of surfaces and syntactic patterns



- We proposed two methods
  - **Ensemble of hidden layers from UTH-BERT**
    - UTH-BERT consists of 12 transformer layers
    - Each layer of BERT captures different features
    - We confirmed effective combinations of the layers for each named entity tag and ensembled them
  - **UTH-BERT with a CRF layer**
    - We used a CRF layer to consider tag sequences



# Results – Subtask1



- Ensemble method outperformed CRF method overall

		Development				FormalRun			
		CR-JA		RR-JA		CR-JA		RR-JA	
	Tag	Ensemble	Crf	Ensemble	Crf	Ensemble	Crf	Ensemble	Crf
F1-score	a	<b>69.61</b>	65.26	<b>100.0</b>	93.62	58.37	<b>58.43</b>	33.58	<b>89.16</b>
	d	<b>80.80</b>	70.88	<b>87.38</b>	82.26	<b>67.05</b>	<b>67.05</b>	7.88	<b>89.40</b>
	m-key	60.61	<b>62.86</b>	-	-	<b>70.63</b>	70.39	-	-
	m-val	<b>0.00</b>	<b>0.00</b>	-	-	<b>65.67</b>	<b>65.67</b>	-	-
	t-key	43.75	<b>46.15</b>	-	-	<b>35.76</b>	35.55	-	-
	t-test	<b>90.91</b>	80.85	<b>100.0</b>	<b>100.0</b>	<b>43.58</b>	43.38	0.00	<b>87.50</b>
	t-val	<b>50.00</b>	42.86	-	-	55.48	<b>55.68</b>	-	-
	timex3	85.00	<b>86.87</b>	<b>100.0</b>	<b>100.0</b>	<b>74.62</b>	74.39	24.49	<b>88.24</b>

- We divided the provided data into training 90% and development 10%.
- Since the result on RR-JA of FormalRun is in stark contrast against Development, the submitted result of Ensemble may include formatting errors.



- Three-stage method

1. Extraction

- We extract nouns or words predicted as named entities by BERT

2. Identification

- We identify named entities tags for the extracted words based on three scores
  - Biomedical dictionary-based score
  - Syntactic pattern-based score
  - BERT-based score

3. Correction

- We apply some rules to merge continuous words into one NE

E.g. brain + metastasis → brain metastasis

<a>

<d>

<d>



- Our method obtained **the best scores** among the participants of the competition

	Tag	CR-JA	RR-JA
F1-score	a	<b>41.52</b>	56.89
	d	<b>41.68</b>	<b>68.45</b>
	m-key	<b>40.00</b>	-
	m-val	22.38	-
	t-key	<b>37.20</b>	-
	t-test	<b>28.17</b>	<b>81.25</b>
	t-val	<b>34.66</b>	-
	timex3	35.02	<b>74.42</b>



- We proposed medical NER methods for few resources
  - Subtask1: **BERT-based approach**
    - Ensemble of hidden layers from UTH-BERT
    - UTH-BERT with a CRF layer
  - Subtask2: **Rule-based approach**
    - Identification based on **weighted three scores**  
(Dictionary, syntactic pattern, and BERT-based scores)
- Results
  - Subtask1: ensemble method obtained good performance
  - Subtask2: our method obtained the best scores among the participants in overall scores