# Enhance Japanese Opinionated Sentence Identification Using Linguistic Features: Experiences of the IISR Group at NTCIR-8 MOAT Task

Guo-Hau Lai, Jyun-Wei Huang, Chia-Pei Gao, Richard Tzong-Han Tsai*
*thtsai@saturn.yzu.edu.tw
Yuan Ze University, Taiwan, R.O.C.

We have described a feature-based system that is designed to recognize opinionated sentences or not. Our system utilizes various features: headlines in newspapers, Japanese sentence patterns, dependency pairs, numeral features and some related to newspapers opinionated words. The experiments show that our feature-based system is feasible and effective.

## System Overview

Our system contains two parts:
1) Generate the features.
2) Recognize opinionated sentences.

Figure 1 shows our system architecture.

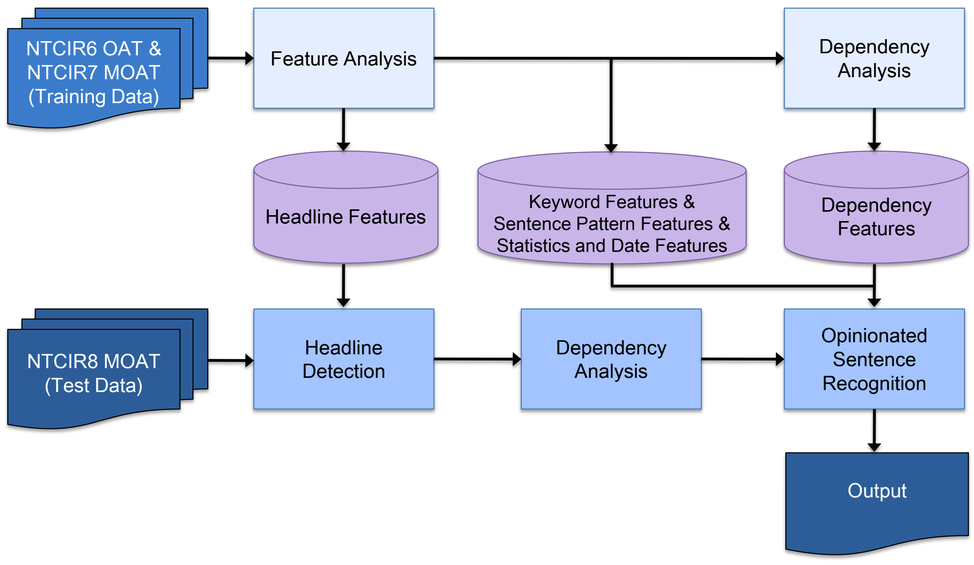


Figure 1: Overall system architecture

## Dataset Analysis

We analyzed NTCIR7 and NTCIR8 sample data (contains 1,601 sentences in total and 499 opinionated sentences) to find useful features for the opinion judgment subtask. We founded these features as follow.

### Headline

We used some special structures and opinion words for opinionated sentence recognition in headlines. 1) A person says something. 2) Opinion words.

### Japanese Sentence Patterns

These following sentence patterns are examples that were frequently used in opinionated sentences.

- について
  この兵士の死とウラン弾の因果関係について、イタリア国防省は否定した。
- そうだ
  パーキンソン病治療や移植医療に貢献しそうだ。

### Keywords

For counting the frequency of each word, we used *MeCab* to perform morphological analysis. Then we compile a opinionated word list and expanded manually by using *Japanese WordNet*.

### Dependency Pairs

We used the *CaboCha* as our Japanese dependency parser, then extracted all dependency pairs and pairs with distances greater than one were not used.

A dependency pair consists of the source element and the sink element. The source element is the named entity (person, organization, etc.) or noun type semantic primitive. However, we only focused on the sink element and extract verbal noun (sahen-noun) according to the pre-compiled list.

### Statistics and Date

The non-opinionated sentences tend to have numbers and these numbers may be statistics, period of time, or a specific time.

- Sentence containing a *statistics data*
  バリ島駐在の日本人向け旅行社によると、バリを訪れる年間約３０万人の日本人観光客のうち６割以上が女性客。
- Sentence containing a *period of time*
  日本でも在日米軍が９５年１２月から９６年１月に沖ノ鳥島の演習で劣化ウラン弾１５２０発を誤射したことが判明。
- Sentence containing a *specific time*
  昨年１０月１２日、インドネシア・バリ島のディスコで爆弾テロが発生、日本人夫婦を含む外國人観光客ら２００人以上が死亡、３００人以上が負傷した。

## Recognition Method

Dataset for training

| | Opinion | Non-opinion |
|---|---|---|
| NTCIR6 OAT | 3,026 | 6,682 |
| NTCIR7 MOAT | 1,562 | 4,323 |
| NTCIR7&8 Sample | 499 | 1,102 |
| **Total** | **5,087** | **12,107** |

### Opinion Score

We extracted the features in the feature sets that was useful for recognizing opinionated sentences. These feature sets consist of a list of patterns or words and we calculate the score for each item by following equation:

$$score(item_i) = \frac{Op(item_i) - On(item_i)}{Op(item_i) + On(item_i)}$$

$score(item_i)$: opinion score of $item_i$
$Op(item_i)$: frequency of $item_i$ appeared in opinionated sentences
$On(item_i)$: frequency of $item_i$ appeared in non-opinionated sentences

We constructed four tables that record opinion score for sentence patterns, keywords, dependency pairs, and numeral features.

### Opinionated Sentence Recognition

1° Our system recognize opinionated sentence according to extracted headlines.
2° For non-headline sentences, the sentence will be analyzed by *CaboCha* for dependency analysis.
3° Our system looks up the features in the dictionaries and sums up the score of each feature appeared in a sentence.
4° If the score calculated by 3° is greater than 0, our systems determines the sentence as opinion.

## Evaluation Result

### Formal Run

We finally submitted three runs as the formal run:
**Run1** Using all features and training without the sample data
**Run2** Delete the headline features and training without the sample data
**Run3** Using all features and all training dataset

The evaluation results are presented in following table:

| | Precision | Recall | F-measure |
|---|---|---|---|
| Run1 | 67.30 | 49.86 | 57.28 |
| Run2 | 67.74 | 47.65 | 55.95 |
| Run3 | 67.86 | 51.53 | 58.58 |

### Effects of feature sets for opinion judgment subtask

We also conducted the experiment to confirm the effects of each feature sets for the opinion judgment subtask. Training data is the same of above and the formal run data is used for the test.

According to the results of formal run, we have confirmed that the headline features are effective. We made comparison between the result of using all feature sets and the result of deleting one feature set (Sentence pattern, keyword, dependency pair, numeral). The following table shows deleted feature sets, each result and difference of F-measure. Larger difference means more effective feature sets.

| Deleted feature | Precision | Recall | F-measure | Difference |
|---|---|---|---|---|
| None | 68.52 | 51.28 | 58.66 | |
| Sentence pattern | 67.81 | 37.04 | 47.91 | 10.75 |
| Keyword | 73.26 | 38.61 | 50.57 | 8.09 |
| Dependency pair | 68.37 | 47.91 | 56.34 | 2.32 |
| Numeral | 66.54 | 53.86 | 59.53 | -0.87 |

The results show numeral feature improve precision but didn't improve recall and F-measure. In some opinionated sentences, these numeral features appear more times than other features and our system output these sentences as non-opinionated sentences.