

Approach

Text Chunking by System Combination

Erik Tjong Kim Sang
University of Antwerp

1. Use a memory-based learner for selecting the best chunk tag for each word.
2. Use a second classifier for boosting the performance of the first.
3. Generate five output streams by training classifiers for five output representations.
4. Combine the five results.

Research questions

- What is the best processing method? (single-pass, double-pass, n-pass)
- What is the best combination method? (we tested nine methods)

CoNLL-2000

1

Tjong Kim Sang

14/09/2000

Results for training data

| double-pass method | O | C | $F_{\beta=1}$ |
|------------------------|--------|--------|---------------|
| Representation | | | |
| IOB1 | 95.49% | 95.84% | 91.59 |
| IOB2 | 95.51% | 95.86% | 91.65 |
| IOE1 | 95.51% | 95.86% | 91.60 |
| IOE2 | 95.64% | 96.03% | 91.97 |
| O+C | 95.67% | 96.12% | 91.97 |
| Simple Voting | | | |
| Majority | 95.88% | 96.24% | 92.34 |
| TotPrecision | 95.88% | 96.24% | 92.34 |
| TagPrecision | 95.88% | 96.24% | 92.34 |
| Precision-Recall | 95.88% | 96.24% | 92.34 |
| Pairwise Voting | | | |
| TagPair | 95.87% | 96.25% | 92.34 |
| Memory-Based | | | |
| Tags | 95.86% | 96.27% | 92.35 |
| Tags + POS | 95.82% | 96.26% | 92.32 |
| Decision Trees | | | |
| Tags | 95.86% | 96.27% | 92.35 |
| Tags + POS | 95.82% | 96.27% | 92.31 |

CoNLL-2000

2

Tjong Kim Sang

14/09/2000

Summary training data results

| training data | O | C | $F_{\beta=1}$ |
|----------------------|--------|--------|---------------|
| sp (MBL, Tags + POS) | 95.80% | 96.28% | 92.40 |
| dp (Majority voting) | 95.88% | 96.24% | 92.34 |
| np (MBL, Tags) | 95.80% | 96.31% | 92.63 |

We chose for the double-pass method in combination with majority voting because the best performing combination methods for the other two require extra processing steps for generating combiner training data.

Majority voting in combination with the n-pass method performs slightly better ($F_{\beta=1}=92.57$) but requires more processing actions.

Majority voting does not perform well in combination the single-pass method ($F_{\beta=1}=91.96$).

CoNLL-2000

3

Results test data

| test data | precision | recall | $F_{\beta=1}$ |
|-----------|-----------|--------|---------------|
| ADJP | 85.25% | 59.36% | 69.99 |
| ADVP | 85.03% | 71.48% | 77.67 |
| CONJP | 42.86% | 33.33% | 37.50 |
| INTJ | 100.00% | 50.00% | 66.67 |
| LST | 0.00% | 0.00% | 0.00 |
| NP | 94.14% | 92.34% | 93.23 |
| PP | 96.45% | 96.59% | 96.52 |
| PRT | 79.49% | 58.49% | 67.39 |
| SBAR | 89.81% | 72.52% | 80.25 |
| VP | 93.97% | 91.35% | 92.64 |
| all | 94.04% | 91.00% | 92.50 |

The large difference between precision and recall rates suggests that there is room for improvement.