A Comparison between Udpipe Tagger and Perceptron-based Tagger

Ali Aljubailan

April 15, 2019

L545

Practical 03: Disambiguation

I tested two tagger on the UD_Finnish-TDT treebank. The first tagger was udpipe tagger. The second tagger was Perceptron-based tagger. After that I evaluated each tagger's performance targetting the row, UPOS. For this evaluation, I utilized the CoNLL-2017 evaluation script. The results of the Perceptron-based tagger and the Udpipe tagger respectively are shown below:

```
149-160-251-62:UD_Finnish-TDT alialjubailan$ cat data.txt
Metrics     | Precision |    Recall |  F1 Score | AligndAcc
------------+-----------+-----------+-----------+-----------
Tokens      |    100.00 |    100.00 |    100.00 |
Sentences   |    100.00 |    100.00 |    100.00 |
Words       |    100.00 |    100.00 |    100.00 |
UPOS        |     90.32 |     90.32 |     90.32 |     90.32
XPOS        |    100.00 |    100.00 |    100.00 |    100.00
Feats       |    100.00 |    100.00 |    100.00 |    100.00
AllTags     |     90.32 |     90.32 |     90.32 |     90.32
Lemmas      |    100.00 |    100.00 |    100.00 |    100.00
UAS         |    100.00 |    100.00 |    100.00 |    100.00
LAS         |    100.00 |    100.00 |    100.00 |    100.00
Alis-MacBook-Air:UD_Finnish-TDT alialjubailan$ python3 conll
Metrics     | Precision |    Recall |  F1 Score | AligndAcc
------------+-----------+-----------+-----------+-----------
Tokens      |    100.00 |    100.00 |    100.00 |
Sentences   |    100.00 |    100.00 |    100.00 |
Words       |    100.00 |    100.00 |    100.00 |
UPOS        |     94.64 |     94.64 |     94.64 |     94.64
XPOS        |     95.81 |     95.81 |     95.81 |     95.81
Feats       |     90.77 |     90.77 |     90.77 |     90.77
AllTags     |     89.75 |     89.75 |     89.75 |     89.75
Lemmas      |     84.52 |     84.52 |     84.52 |     84.52
UAS         |    100.00 |    100.00 |    100.00 |    100.00
LAS         |    100.00 |    100.00 |    100.00 |    100.00
Alis-MacBook-Air:UD_Finnish-TDT alialjubailan$
```

The results of evaluating the Perceptron-based taggers (the above table) and the Udpipe tagger

As can be seen, udpipe tagger achieved 94.64/100 points for Precision, Recall, F1 Score, AligndAcc coloumns for UPOS. On the other hand, Perceptron-based tagger scored 90.32/100 point for the same coloumns. Probably the difference does not seem to be high, but I think that although it seems so (almost 4 points higher for udpide), this trivial difference can make a considerable difference in efficiency.

However, this difference might be explained by that the Perceptron-based tagger is more simply written than the Udpipe one, which is one of the known models today. If we consider this, I think that somewhat can give a good indication about the efficiency of the Perceptron-based tagger, as it is more simply prepared and treated and in spite of that it can show a fair competition to the Udpipe tagger.

Furthermore, there might be other certain factors that could play a role in the result between the two tagger. For instance, it can be noticed that the row Lemmas sharply differs from the Udpipe to the the Perceptron-based tagger, by above 15%. That is, whereas Udpipe scored in Lemmas 100/100, the Perceptron-based tagger has only scored 84.52/100. Another example is the row Feats, in which Udpipe scored 100/100, by more than 10 point above the score of the the Perceptron-based tagger. I think that probably because the achievement of UPOS is likely to be affected by the achievement of other rows such as Lemmas and Feats in prarticular, this could partially explain the differences of scoring for the both taggers.

In sum, the comparison the was made between the Udpipe tagger and the Perceptron-based tagger on the UD_Finnish-TDT treebank resulted in observably higher socres for the Udpipe tagger. On the row UPOS in particular and all other rows in general, the Udpipe tagger enjoys a higher score and thus more efficiency rate than the simple Perceptron-based tagger. However, the Perceptron-based tagger seems to be able to achieve a higher score and enhance its performance if it is promoted a little bit more.