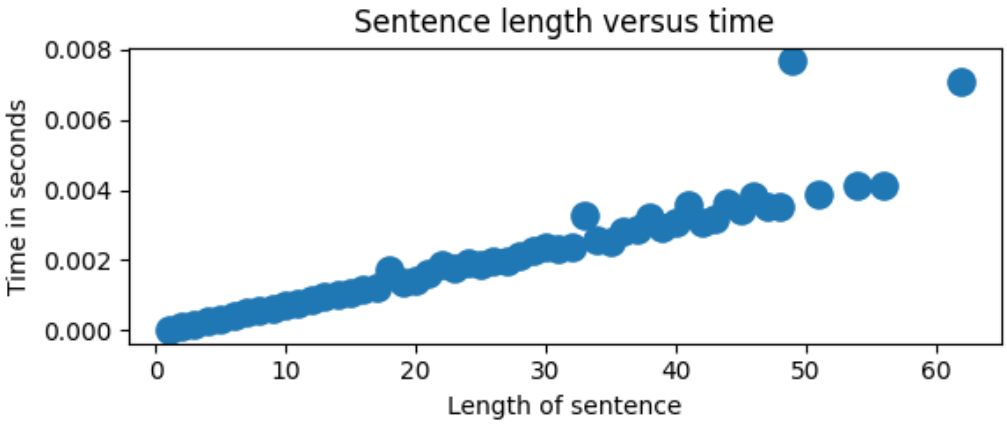
Discussion

The graph (eval-t-l) shows that the implemented algorithm is very fast and Viterbi does not take longer than 0.015 seconds. With increasing sentence length, the time in seconds that the algorithms need to calculate the POS taggings for the sentence, also increases linearly. The total time for the evaluation test including the HMM training takes no more than 5 seconds.



The table (accuracy_score) however shows that accuracy is only 79% and could be improved, e.g. by unknown word handling, smoothing or a larger corpus. While the precision is very high for conjunctions and numbers, it is comparatively low for pronouns, verbs and determiners and unknown words. On the other hand the table displays high recall for determiners, verbs and especially punctuation, whereas the recall for adverbs and adjectives is quite low and extremly low for unknown words (0,09). Consequently punctuation in f1 achieves the highest result, followed by conjunctions and adp. Unknown words (X) reach by far the lowest score.

```
Comparing gold file "de-eval.tt" and system file "dd-test.t"

Precision, recall, and F1 score:

  DET 0.6867 0.8939 0.7767
 NOUN 0.8331 0.7037 0.7629
 VERB 0.6977 0.8630 0.7716
  ADP 0.7807 0.8740 0.8247
    . 0.9503 0.9328 0.9415
 CONJ 0.8773 0.7887 0.8306
 PRON 0.6451 0.7230 0.6819
  ADV 0.8604 0.6811 0.7604
  ADJ 0.7528 0.6504 0.6978
  NUM 0.9450 0.7000 0.8043
  PRT 0.8119 0.8013 0.8066
    X 0.6667 0.0909 0.1600

Accuracy: 0.7914
```