

Meine einzigen Änderungen am Code waren eigentlich nur die Vokabulargröße zu verkleinern und die Anzahl Hidden Layers zu verringern. Das aus praktischen Gründen, da die GPU's ständig ausgelastet waren und ich nicht ewig rumrechnen und den Platz belegen wollte und die Zeit drängte. Zunächst auf 40'000 respektive 512, aber als sich abzeichnete, dass auch das zu lange dauern würde bin ich hinab auf 30'000 fürs Vokabular und 265 Layers. Die Perplexität des dadurch trainierten Systems betrug 4.61 nach 6 Epochen des Trainierens. Ich hätte gerne ein paar der vorgeschlagenen Ideen implementiert, aber leider fehlte mir auch hierfür die Zeit bzw. GPU's. Ich habe ein wenig herumexperimentiert aber letztendlich den funktionierenden Code so gelassen wie er ist. Probleme bereitet hat das Übersetzen, denn der Befehl der im daikon readme steht, funktioniert (bei mir zumindest) nicht und produzierte stur kryptische Fehlermeldungen. Ich habe mir lange den Kopf darüber zerbrochen, bis ich merkte, dass das echo das Problem war. Und schlussendlich habe ich den Befehl folgendermassen abgeändert:

```
CUDA_VISIBLE_DEVICES=3 daikon translate -i corpus.dev.de > eng_out.txt
```

Ein anderes Problem war der Bug mit dem Vokabular aber das wurde ja im Nachhinein gefixt. Der übersetzte Text findet sich in eng_out.txt (der model folder und die trainings und dev files sind in meinem local repository, aber irgendwie schaffe ich es nicht sie zu pushen. Da aber in der Aufgabe ja nur der Code, der übersetzte Text und der Bericht verlangt sind, habe ich einfach letztere heruntergeladen und auf github ins remote repository hochgeladen, ich hoffe das ist in Ordnung).

Perplexität nach 6 Epochen:

```
2018-05-19 15:49:36,426 - DEBUG - Epoch=6, iteration=88000/91202, samples/second=148.17
2018-05-19 15:51:50,943 - DEBUG - Epoch=6, iteration=89000/91202, samples/second=148.68
2018-05-19 15:54:06,139 - DEBUG - Epoch=6, iteration=90000/91202, samples/second=147.93
2018-05-19 15:56:20,776 - DEBUG - Epoch=6, iteration=91000/91202, samples/second=148.55
2018-05-19 15:56:47,766 - DEBUG - Epoch=6, iteration=91202/91202, samples/second=741.02
2018-05-19 15:56:47,993 - INFO - Perplexity on training data after epoch 6: 4.61
```