Evaluation

The github repository: [Link](https://github.com/Laura-Christoph/mt-exercise-03)

The Evaluation etc are both in here and all the data etc are in the repository

# My data

«I led Babieca into as much shade as I could devise, tied him to a bush,

and crawled under it with my unlucky brains. While taking refuge here

I had a fall in fortune. You will conceive, O admirable reader! that

the sun, this false friend in whom I had reposed my trust, having dealt

with me in this false spirit, there was no longer that poetry in my

temper with which I had begun my journey. I was beset with doubts. If a

face so bright, so open, so intelligible could hide such malice, where

was the candour of the world? By this pertinent reflection my thoughts

were carried to the poor woman who had also shared my trust. Perchance

it was not the part of wisdom to bestow the tenth portion of my

inheritance upon a beggar in the road. Sorely considering this aspect

of the case I took forth my pouch, and pouring my little means into my

hand, not without a pang that one palm could hold it all, behold! in

lieu of nine crowns I discovered that I had but eight.»

This is a snippet of my data from the book Fortune.

# Expectations

From my Model, I except a similar narrative style and tone, which seems to be reflective and somewhat introspective. I also expect the style to be rather old.

# Sample

“by his own hand upon the road . <eos> Very addressed upon them , he said in the <unk> ,

while even we could be set stick with the aid of princes . <eos> At his veins it was that

another of the <unk> spoke such an strange days room was not men , until I heard my native mountains

and began to take the <unk> of the Castilian ; and then very grave , it was all more fortunate

than an lawful addresses . <eos> But owing to a cord was impossible . <eos> For no sooner had it”

I find this to be a very confusing snippet, the sentences don’t make much sense, however my data wasn’t very big anyways.

Valid perplexity

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Valid perplexity** | Dropout 0 | Dropout 0.2 | Dropout 0.4 | Dropout 0.6 | Dropout 0.8 |
| **Epoch 1** | 269.79 | 241.33 | 250.85 | 218.21 | 242.23 |
| **Epoch 2** | 145.02 | 147.01 | 146.08 | 152.47 | 183.88 |
| **Epoch 3** | 121.12 | 116.14 | 119.45 | 127.32 | 162.94 |
| **Epoch 4** | 104.89 | 106.88 | 104.49 | 115.51 | 153.22 |
| **Epoch 5** | 100.69 | 96.72 | 98.14 | 105.44 | 146.96 |
| **Epoch 6** | 96.05 | 90.53 | 92.39 | 99.5 | 139.49 |
| **Epoch 7** | 95.23 | 87.43 | 90.49 | 94.61 | 136.29 |
| **Epoch 8** | 100.46 | 86.69 | 86.12 | 93.23 | 133.36 |
| **Epoch 9** | 88.83 | 86.17 | 84.7 | 90.38 | 130.35 |
| **Epoch 10** | 90.77 | 87.05 | 82.76 | 88.58 | 128.34 |
| **Epoch 11** | 89.24 | 79.52 | 81.15 | 86.33 | 127.08 |
| **Epoch 12** | 88.4 | 78.85 | 80.05 | 85.79 | 125.15 |
| **Epoch 13** | 88.33 | 78.87 | 79.18 | 83.36 | 123.82 |
| **Epoch 14** | 88.39 | 78.12 | 77.83 | 83.09 | 122.22 |
| **Epoch 15** | 88.17 | 78.52 | 78.15 | 82.92 | 121.4 |
| **Epoch 16** | 88.09 | 78.15 | 73.92 | 81.36 | 120.8 |
| **Epoch 17** | 88.06 | 78.09 | 73.49 | 81.22 | 118.2 |
| **Epoch 18** | 88.06 | 78.05 | 73.28 | 82.07 | 117.73 |
| **Epoch 19** | 88.07 | 78.03 | 73.04 | 77.09 | 119.05 |
| **Epoch 20** | 88.05 | 77.99 | 73 | 76.69 | 113.6 |
| **Epoch 21** | 88.04 | 77.97 | 73.25 | 76.24 | 112.94 |
| **Epoch 22** | 88.03 | 77.98 | 72.5 | 76.15 | 111.97 |
| **Epoch 23** | 88.03 | 77.98 | 72.51 | 75.69 | 111.34 |
| **Epoch 24** | 88.03 | 77.97 | 72.22 | 75.69 | 110.67 |
| **Epoch 25** | 88.03 | 77.97 | 72.12 | 75.21 | 109.89 |
| **Epoch 26** | 88.03 | 77.97 | 72.03 | 75.04 | 109.34 |
| **Epoch 27** | 88.03 | 77.97 | 72 | 74.93 | 108.71 |
| **Epoch 28** | 88.03 | 77.97 | 72.03 | 74.92 | 108.55 |
| **Epoch 29** | 88.03 | 77.97 | 72.02 | 74.3 | 107.94 |
| **Epoch 30** | 88.03 | 77.97 | 72.02 | 74.45 | 107.16 |
| **Epoch 31** | 88.03 | 77.97 | 72.01 | 74.08 | 107.11 |
| **Epoch 32** | 88.03 | 77.97 | 72.01 | 73.89 | 106.59 |
| **Epoch 33** | 88.03 | 77.97 | 72.01 | 73.91 | 105.63 |
| **Epoch 34** | 88.03 | 77.97 | 72.01 | 73.76 | 105.58 |
| **Epoch 35** | 88.03 | 77.97 | 72.01 | 73.7 | 104.97 |
| **Epoch 36** | 88.03 | 77.97 | 72.01 | 73.67 | 104.79 |
| **Epoch 37** | 88.03 | 77.97 | 72.01 | 73.65 | 104.43 |
| **Epoch 38** | 88.03 | 77.97 | 72.01 | 73.61 | 104.19 |
| **Epoch 39** | 88.03 | 77.97 | 72.01 | 73.59 | 103.93 |
| **Epoch 40** | 88.03 | 77.97 | 72.01 | 73.62 | 103.25 |

# Evaluation

Looking at all the numbers and graphs, the model with 0.4 dropout performed the best, closely followed by the model with 0.6 dropout. The model with 0.8 dropout performed the worst.

## Sample Model 0.4

“by his own hand . <eos> Being a information that addressed upon them , and our noble courtier were ,

while even out of all the night was the aid of princes board . <eos> Yet the duke was again

thrown with the <unk> manner , and with unhappy room , the duke , until his favourite material Nothing creep

into the wood , the Count of Nullepart was broken with the mob . <eos> The <unk> of onlookers ,

<unk> powers , and with a <unk> spot , a cord was in the finest and arrogance of his ears”

## Sample Model 0.8

“by her imagination that Sir Richard Pendragon that being that addressed upon them , and our noble man ! <eos>

The even <unk> mankind , in stick of the aid in princes board at this <unk> of governance . <eos>

“ And I <unk> , ladyship , ” said the sentinel , men , until which I felt her creep

and began to take the <unk> of the Castilian ; and the very way , it think all could never

met an lawful addresses and so and his state , me was went . <eos> For no accomplished of it”

## Sample Evaluation

Model 0.4 seems to produce text with slightly more coherence in terms of grammar and syntax compared to Model 0.8. The sentences are structured more conventionally, although some phrases still lack clarity and contain unidentified tokens ("<unk>").

Model 0.8 appears to produce text that is more imaginative and includes more diverse vocabulary. However, this comes at the cost of coherence and clarity, with some sentences being less grammatically correct and containing phrases that may be difficult to interpret.