DDA5001 PROJECT PART I: PRE-TRAINING A GPT MODEL ON SHAKESPEARE DATASET

Author
PENG Qiheng
225040065
The Chinese University of Hong Kong, Shenzhen
School of Data Science

1 Result

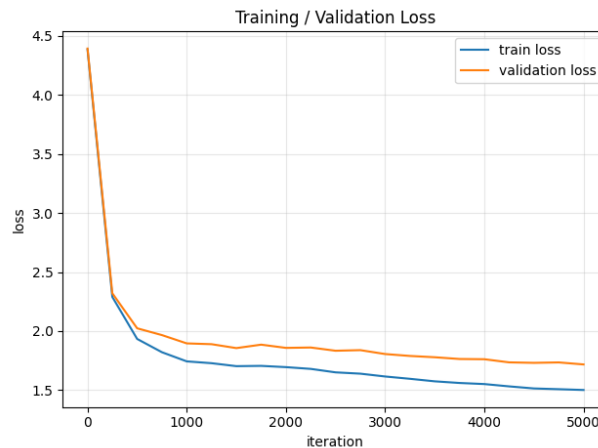The the training loss and validation loss are shown in Figure 1.

Figure 1: Training and Validation Loss

2 Difficulties and Observations

Difficulties:
(a) Dependency conflict issue while running pip install.
—- Solution: Install the corresponding pytorch and other packages based on my CUDA version.

Observations:
(a) The training loss is lower than the validation loss after 1000 iterations, which may indicate slight overfitting.
(b) The generated text is still quite noisy, which may be due to the small size of the Shakespeare dataset.

3 Appendix

3.1 Generated Text (One Sample)

Were all the with well him woman be distill:
Wherink when sing feel'd with the pity
To that which he watter in the with strewar.

Second Servingman:
He in here sone! which or are he ward there a given we hate follower of the mon
a cred to-dear, whom of my cannot their: speak be somession; and a game by the sin
The seast just that theer-gent.

CORIOLANUS:
Tell me.
If not all man:
I did the kill stander therefore the is long:
And away tearshife good him him.

SLY:
'Tis dready and deman, the that a