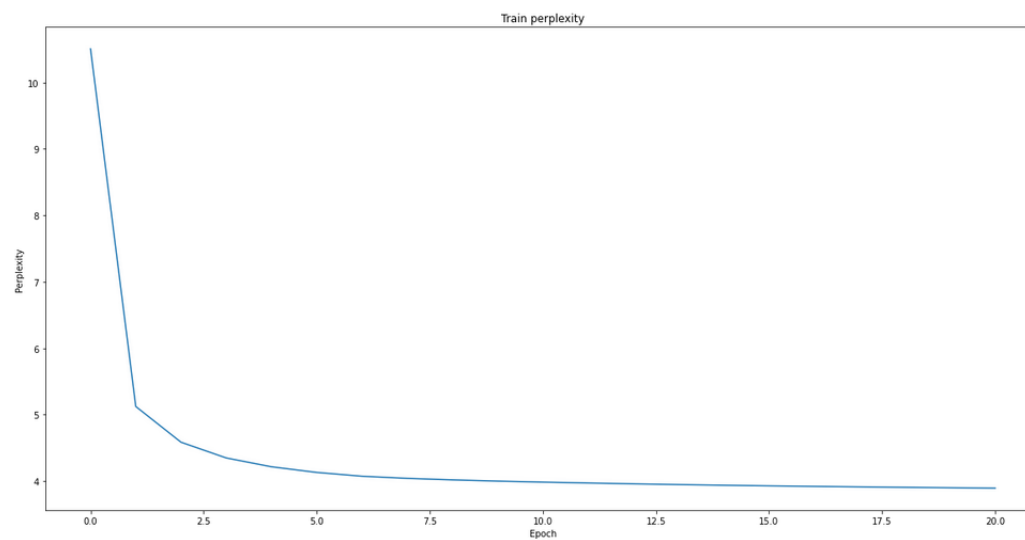
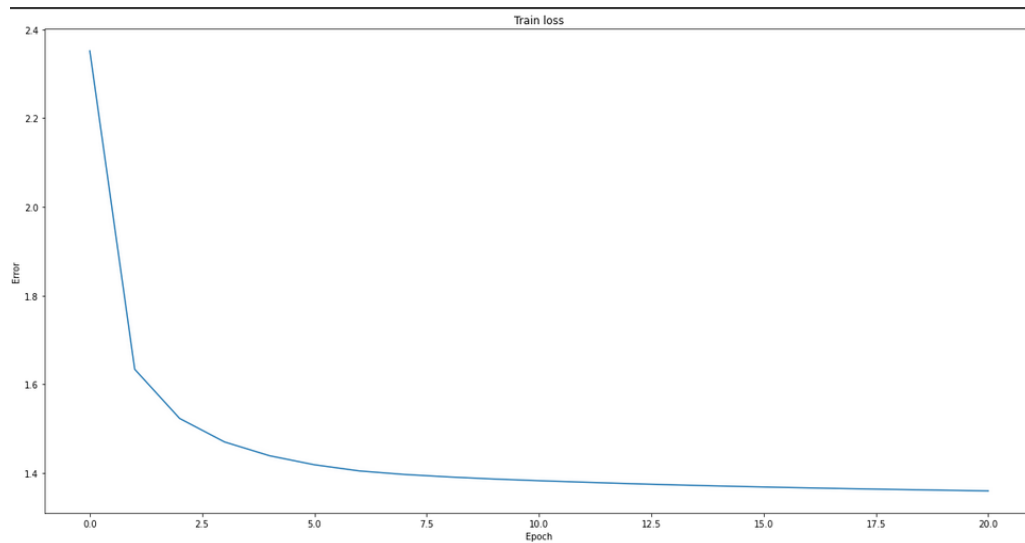


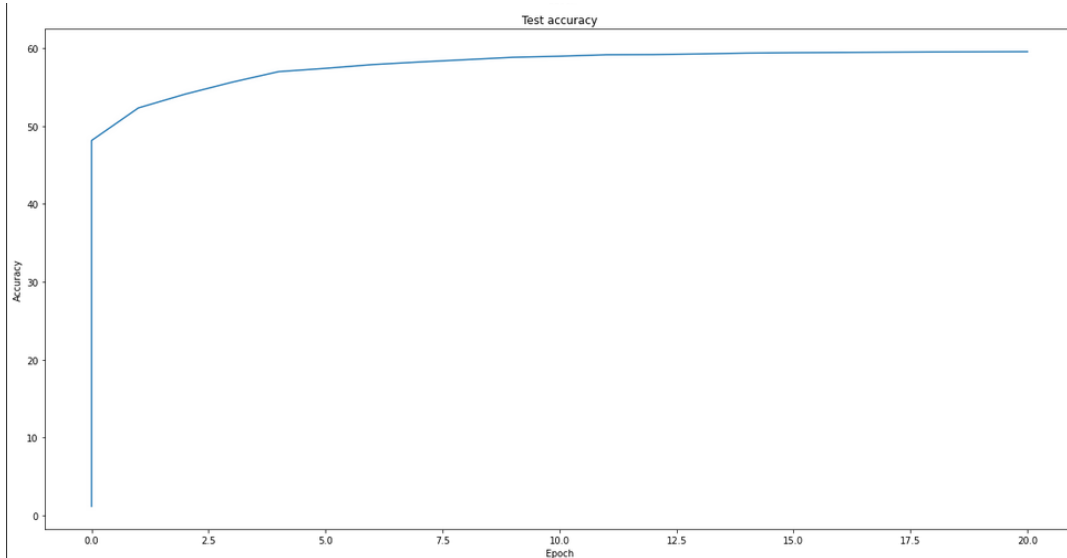
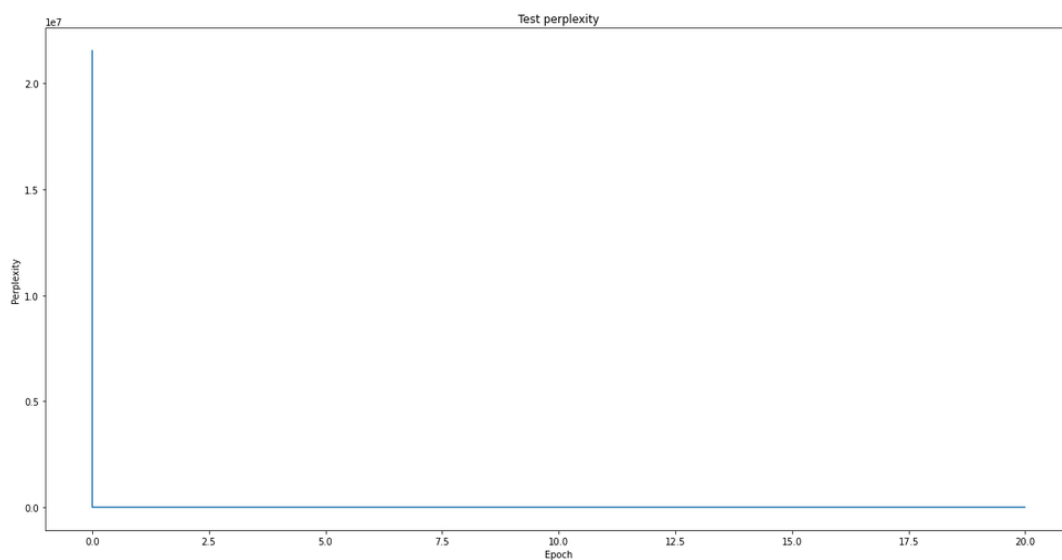
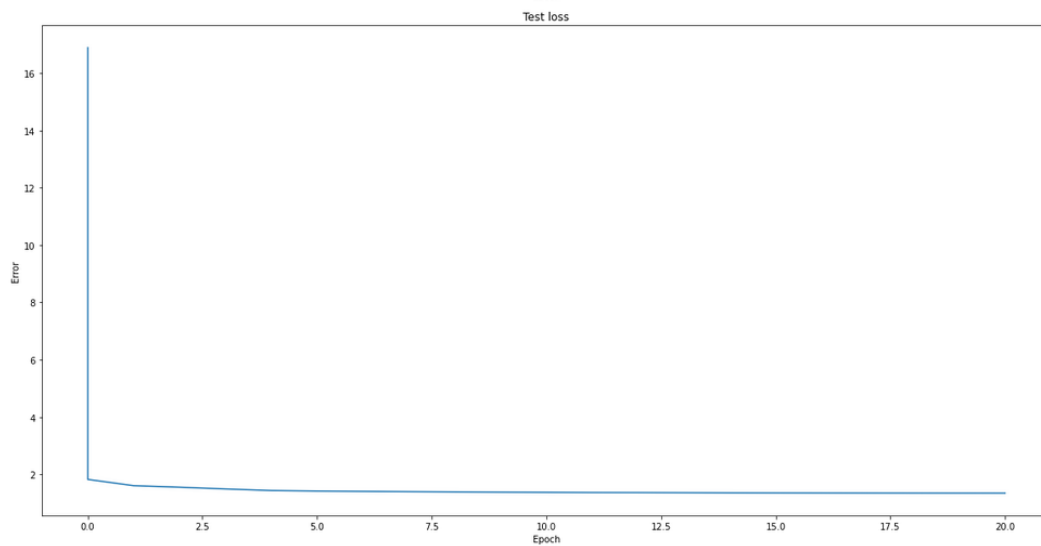
Part 9 Short Answer Questions

1. Plots for GRU training on Harry Potter Dataset

Hyperparameters:

- ❖ SEQUENCE_LENGTH = 50
- ❖ BATCH_SIZE = 256
- ❖ FEATURE_SIZE = 512
- ❖ TEST_BATCH_SIZE = 256
- ❖ EPOCHS = 20
- ❖ LEARNING_RATE = 0.001
- ❖ WEIGHT_DECAY = 0.0005





2. Out final **test accuracy** = 60%, and our final **test perplexity** = 3.83205035239, as our final test loss = 1.3434

3. Favorite sentence with Prompt: 'Harry Potter and the'
generated with max

- Harry Potter and the stone stone stone stone of the castle and said to his feet. "What do you think you to the one of th

generated with sample

- Harry Potter and the owlbrewing Haw had Deer Rourstakene house should knows by Ginny on to," "You ang Permione's hoo, in

generated with beam

- Harry Potter and the first time," said Hagrid, and Hermione looked at Harry," said Harry, looking at Harry. Harry looked

4. The sampling method that generated the best results was beam sampling. The reason I chose beam sampling to be the best was because the outputs generated with beam sampling generally only included words that exist in the text and were the most understandable overall. I believe that beam sampling performed the best because it was able to keep some variety in the output while also generating output that used characters that were likely to appear.

5. I believe the outputs are produced when TEMPERATURE = 2, as it had the funniest results due to more variety in the generated output. Having the TEMPERATURE = 0 makes the model very confident in the prediction, the TEMPERATURE in the range (0, 1) will make the model feel confident in its own prediction but still use some randomness, the TEMPERATURE = 1 means the model directly applies F.softmax() without any change, and the TEMPERATURE > 1 leads to much more random outputs. Negative TEMPERATURE would lead to the outputs being generated with the least likely characters.

New Corpus

1. The corpus we chose was *The Adventures of Sherlock Holmes* (<https://norvig.com/big.txt>). After removing the end of that file, which seemed to just be a glossary containing words in a random order, our new corpus contained 6,344,451 characters.
2. The main difference I noticed between the sentences generated by the Harry Potter text and the Sherlock Holmes text was that common words and type of words that ended up being chosen were totally different. The outputs generated by the Harry Potter text generally included at least one attempt to write a name down, and also would have many outputs include quotes. On the other hand, the outputs generated by the Sherlock Holmes text are more attempts at recreating a text similar to an excerpt from the text. Outside of the seed words, the outputs did not include names and focused more on common words seen in the text.
3. For the outputs of the sampling method on the Sherlock corpus we used
TEMPERATURE = 2

generated with max

1. Sherlock Holmes and the state of the state of the state of the state of the state of the state of the state of

generated with sample

1. Sherlock Holmes and the exe changed. I take it, and that they would use, are ween the rule under state diments, on the laxt
2. Sherlock Holmes and the production to you. Monstitutif the gold oflicteep our prelius for the modernment.. The Narid of a bl
3. Sherlock Holmes and the fring their purstication on the organs of a man interfue, the tissue of unkeptn oliniage pounsing g
4. Sherlock Holmes and the plants vew one, of manter ausounced the involfular from them, ladid tounhed, with the muscles inject
5. Sherlock Holmes and the room said, and a tamen-shouldgeneren fither colled the puls, those when lifes Prompectale, likning,
6. Sherlock Holmes and the nuture, the owen had been considered. It is applied up. The respoons of the stain of dovaling his le
7. Sherlock Holmes and the colonies of the windockin meen over months in the dision byoness who right." "That has adravence, Af
8. Sherlock Holmes and the medumary the arilt of the heats; or lefty meduma would have not sure then usbut does tarked to stop

9. Sherlock Holmes and the same aN hose ownured it will, but I don't horse?" I was " said Pierre slave ogen of the farmers with
10. Sherlock Holmes and the manunarday and naemial, mulful an upock, and lip-essenities lightened their charan, This was commad;

generated with beam

1. Sherlock Holmes and the muscles when there was not that there would not have been considered that there was nothing to their
2. Sherlock Holmes and the muscle of their constitution of their constitution of their constitutions of their constitution of th
3. Sherlock Holmes and the muscles when there is not only that there would not take that there was not that there would have the
4. Sherlock Holmes and the muscle of the muscles, and there were there was not that that there was nothing, and there was their
5. Sherlock Holmes and the muscles of the muscles and their constitution of their constitution of their constitution of their co
6. Sherlock Holmes and the muscles of their number of their constitution, and their constitution of their constitution of their
7. Sherlock Holmes and the muscles of the muscles of their constitutions of their constitution, and their constitutions of their
8. Sherlock Holmes and the muscle of their constitution of their constitution of their constitutions of their constitution of th
9. Sherlock Holmes and the muscles of their government and their constitutions of their muscles which there was not only busines
10. Sherlock Holmes and the muscles of their muscles, and there is usually through their constitutions of their princess, and the

New Architecture

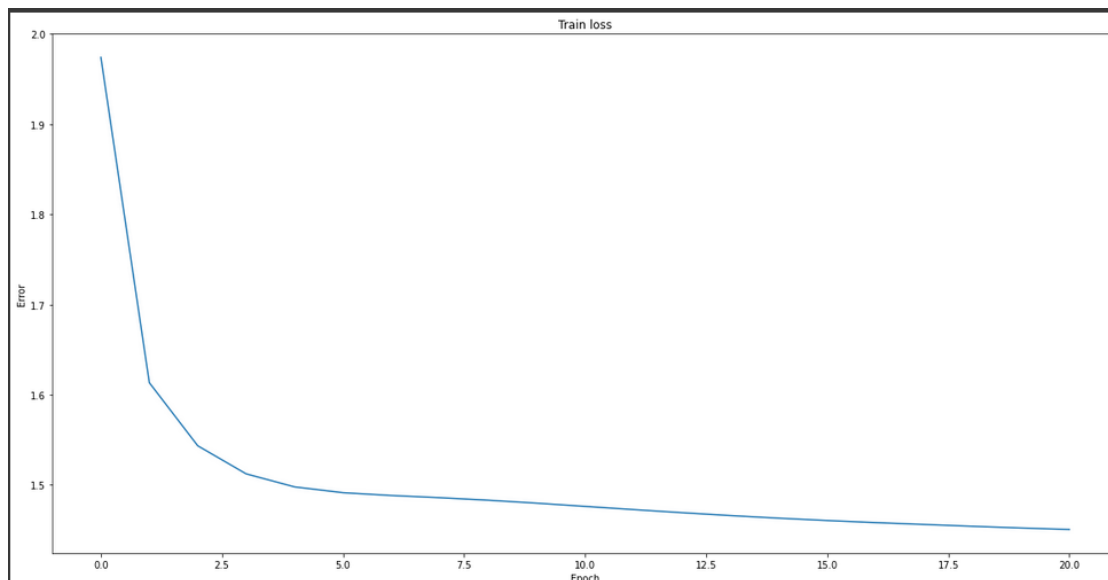
1. The design of the better architecture occurred by adding a fully connected layer (`nn.Linear(self.feature_size, self.feature_size)`) after the GRU layer, but before the decoding layer.

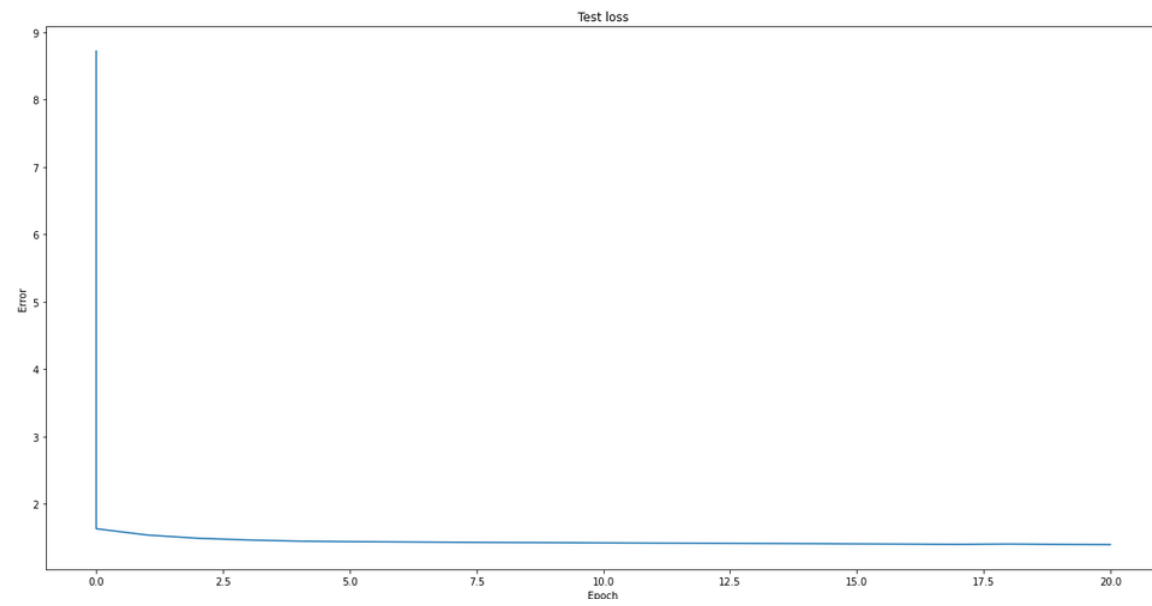
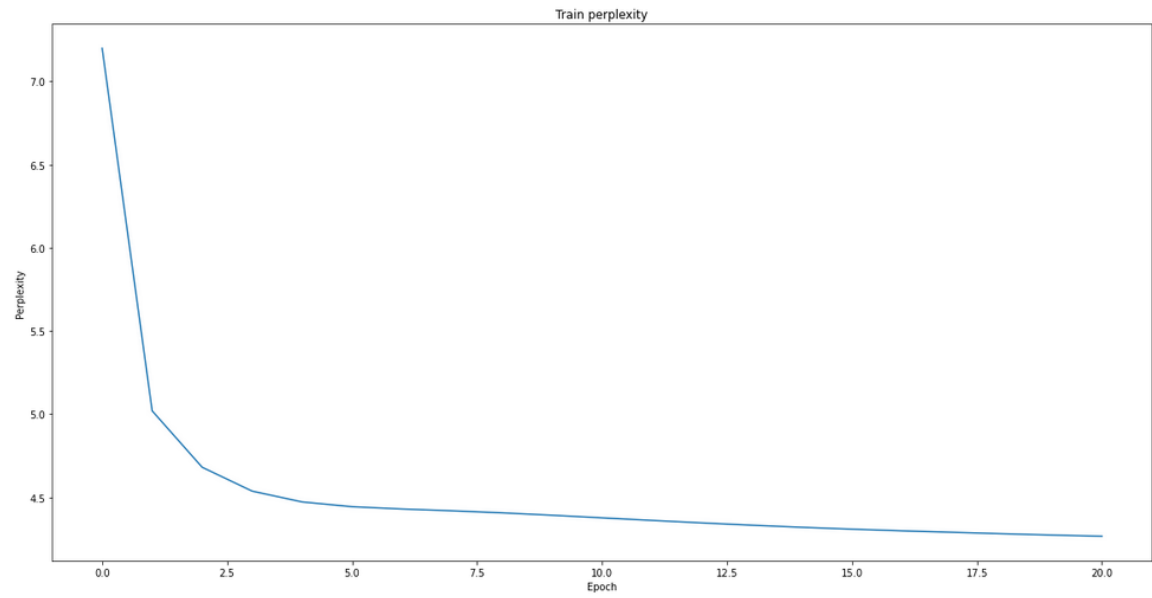
The final design was

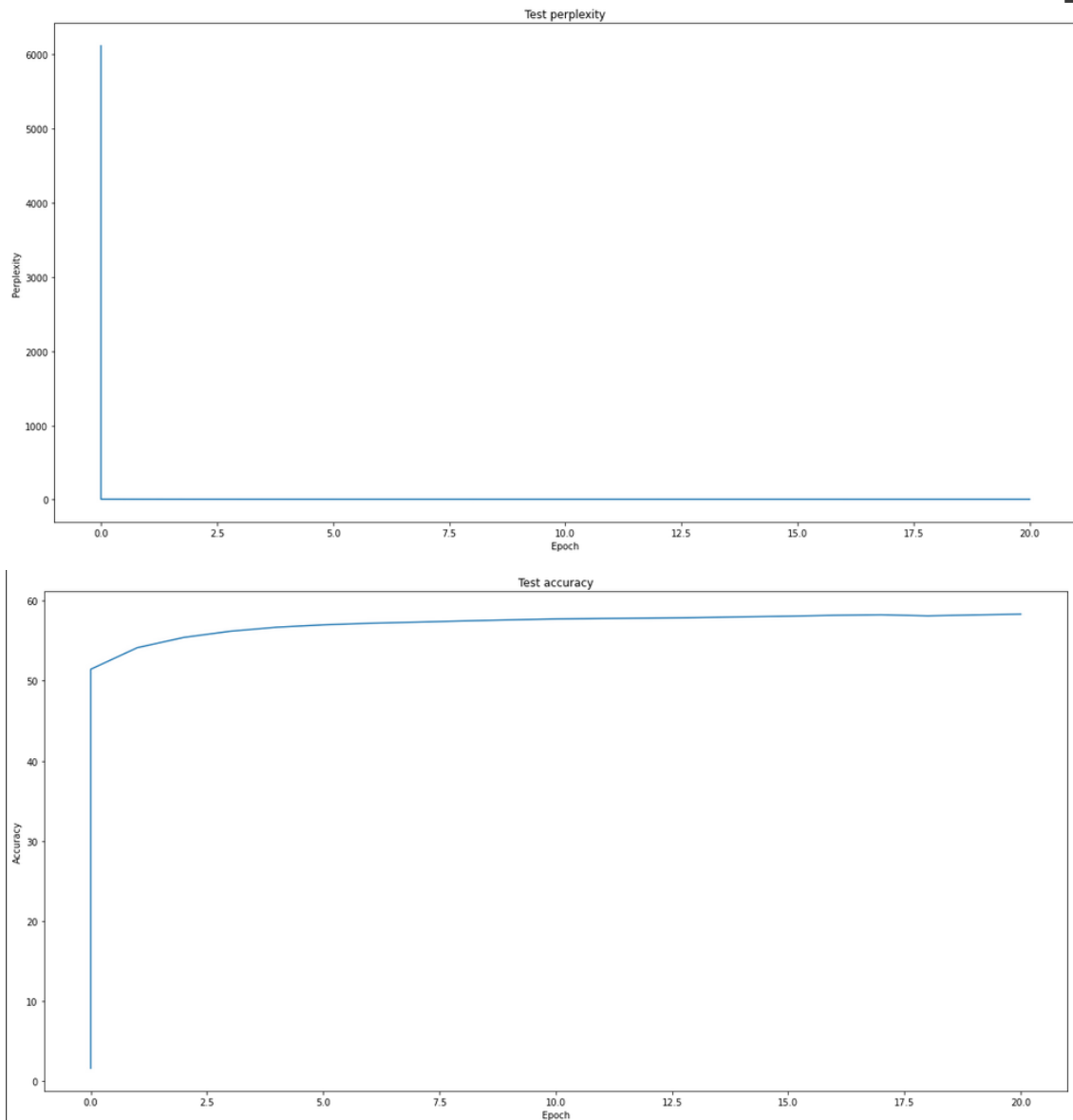
- a. encoding layer (`nn.Embedding(self.vocab_size, self.feature_size)`)
- b. GRU layer (`nn.GRU(self.feature_size, self.feature_size, batch_first=True)`)
- c. Linear layer (`nn.Linear(self.feature_size, self.feature_size)`)
- d. ReLU activation function
- e. decoding layer (`nn.Linear(self.feature_size, self.vocab_size)`)

We also tried to stack several encoders on top of each other but the parameter numbers did not match well so the network was unable to run.

2. Our lowest **test perplexity** = **3.64917093**, as our lowest test loss = 1.2945







3. For the outputs of the sampling method on the Harry Potter corpus we used
 TEMPERATURE = 2

generated with max

1. Harry Potter and the street street and stared at the stairs and staring at the stairs and staring at the stairs and star

generated with sample

1. Harry Potter and the entmus changed inches were far that Mr.R Crong and he had eats Water. "She"ll unay bad a cernent.
2. Harry Potter and the second." The two week and looked line Hagrid raised him slightened. "Do you stay," said Lavin. Harr

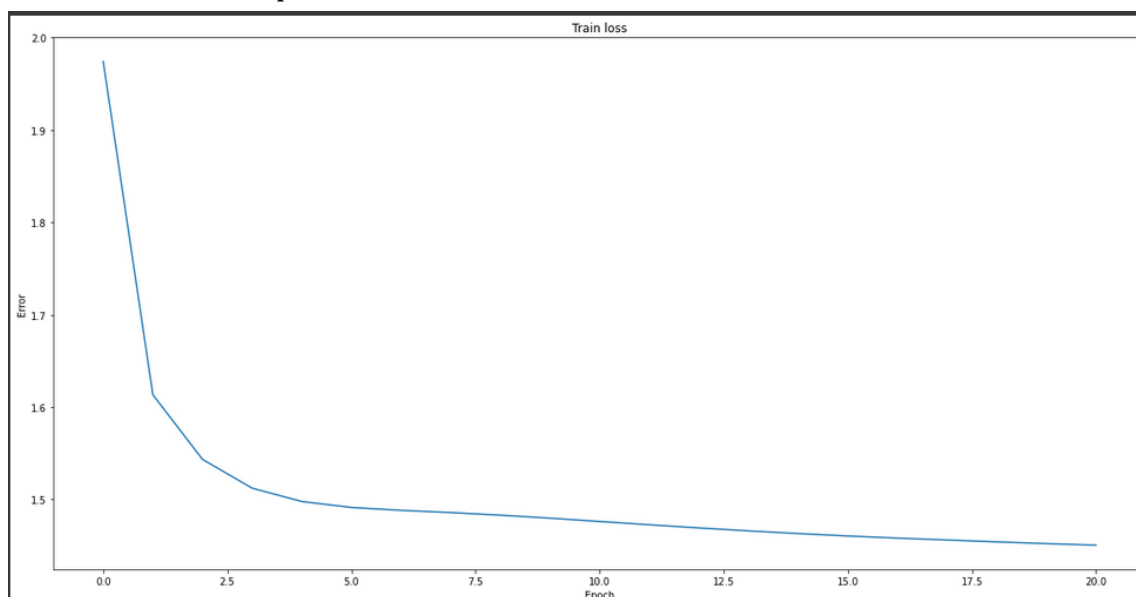
3. Harry Potter and the smited floor. "Cale you are that we sorry, Harry -- and this acid Uon." "Syed," said Harry grignly
4. Harry Potter and the of her arress the cacked, gnimper he had arrixed nearly and a rules uncle lost A laugh sneer. "Ohap
5. Harry Potter and they set off to sage the floor closer who was standing around the boggy for to year a chair that appeare
6. Harry Potter and the distant; People she had sprinted. Tow," she might be with a smiling golden, which severed pharing a
7. Harry Potter and the onlu wall foolish blooked. Ron, despintang. "I's really to help up and Snape line ah you who had to
8. Harry Potter and the moment, looking and turned Ron, whice she had nothing. Rupo I see the tine; has sadeway, Righang th
9. Harry Potter and these next the Hug's grow whose wall after alf of studenot. Theys come much conscept with and lengers a
10. Harry Potter and the tower Appeared and quickly he followed him his clutchet, all shapping a sit of rawn. Nape stood exv

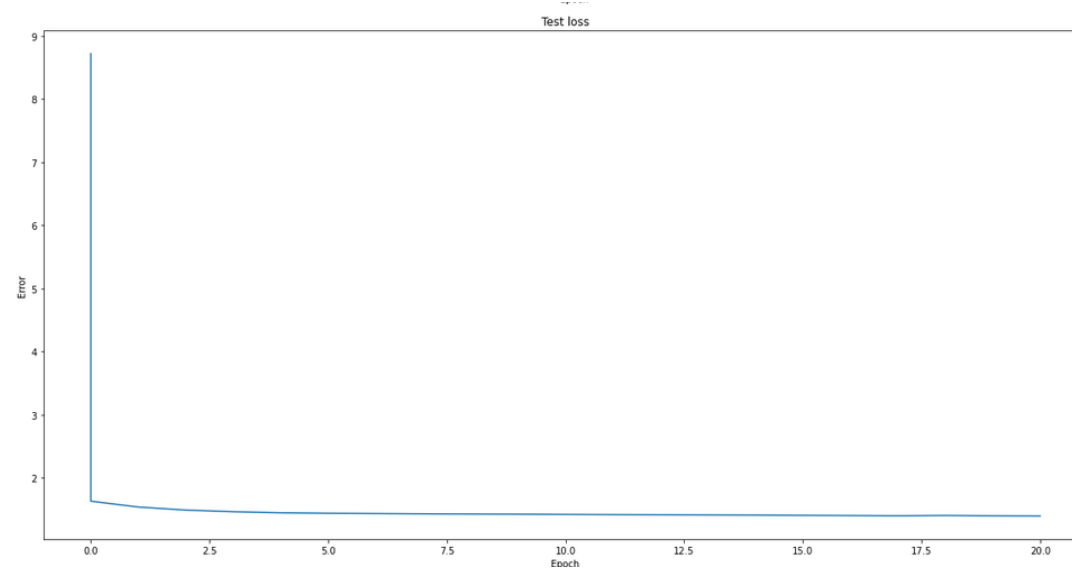
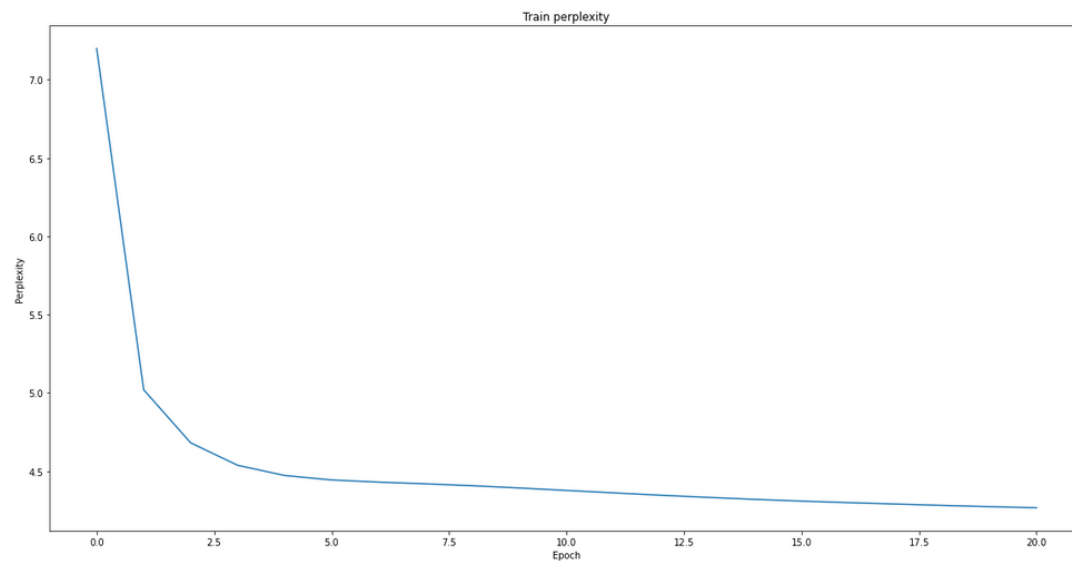
generated with beam

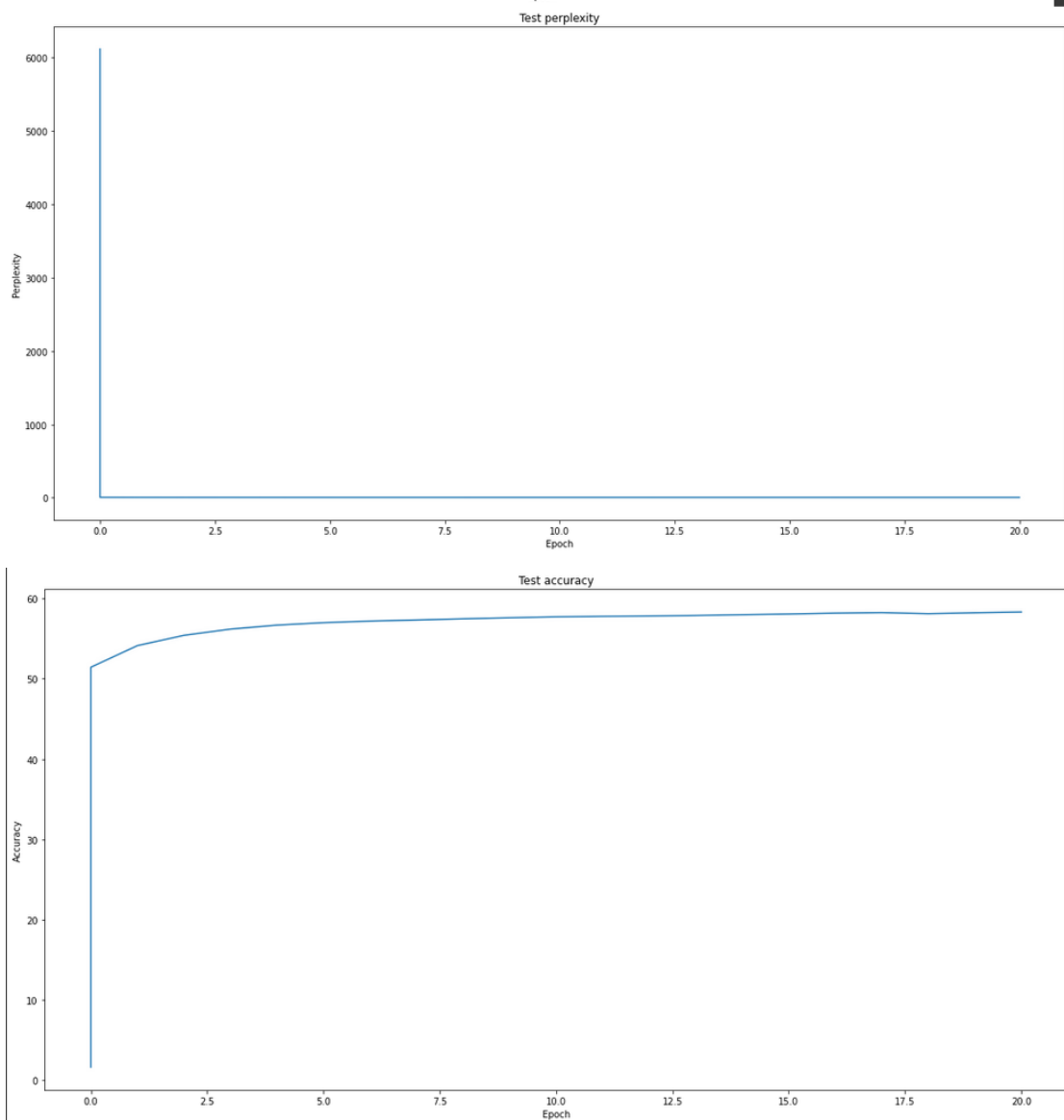
1. Harry Potter and they wouldn't have been there," said Hermione, looking as though he would have been looking around with
2. Harry Potter and the Ministry," said Hermione. Harry looked around. Harry was looking around him. Harry had been looking
3. Harry Potter and they would have never been looking around," said Dumbledore. Hermione, looking around. Harry looked into
4. Harry Potter and they wouldn't have been looking," said Hermione, looking at Harry," said Hermione, which had been lookin
5. Harry Potter and they wouldn't have been here," said Hermione, looking loudly. Harry would have been talking into the air
6. Harry Potter and they would have thought he wouldn't have been going," said Ron, looking around. Harry, who had been been
7. Harry Potter and they wouldn't have gone," said Dumbledore, looking around. Harry looked at Harry, looking around. Harry
8. Harry Potter and they would have been very little," said Dumbledore," said Dumbledore. Harry looked as though he had been
9. Harry Potter and they would have been," said Hermione, looking around. Harry looked around. There was looking through the
10. Harry Potter and they would have been trying to have been," said Hermione, looking down into the door. Harry and Hermione

LSTM Questions

1. The main difficulty that came when trying to train the model was to identify that None was different from a tuple of (None, None). For the initial input, both the hidden and cell layers are initialized to None which causes an error when trying to apply the LSTM layer with a tuple of (None, None) as input. After looking at the code for LSTM, we realized that instead we needed to just send in a None tuple as input to the LSTM layer for it to handle the case by itself.
2. The results of the LSTM were very similar to GRU, but did seem to be very marginally better when looking at Test Accuracy. However, the LSTM did seem to converge faster than the GRU. In our code we ran the LSTM for both the Harry Potter text as well as our chosen Sherlock Holmes text to compare the results, and have included the plots for the LSTM on the Sherlock Holmes text below.







3. For the outputs of the sampling method on the new corpus we used

TEMPERATURE = 2

generated with max

1. Sherlock Holmes and the states of the states of the states of the states of the states of the states of the st

generated with sample

1. Sherlock Holmes and the cases of the Rossor surface membrous forms of whub. N36. 79C...-herform Reonstruck, and in _ Rosto
2. Sherlock Holmes and the epi-med-uart. |. But myself. I. The Roosemer of Macoma, appared by the stret mary swokling. I gave y

3. Sherlock Holmes and the Ule had enmerace. It had the his native of those vois, that the settle of EBhigo vay, they revel he
4. Sherlock Holmes and the deaterial parests._--bly must be considered to Nell that people as a hingly even everyheldge or intr
5. Sherlock Holmes and the inclus of this crops my bring, hould be one a miles of drar vague, but where unever cions, own of qu
6. Sherlock Holmes and the hyor and musita as to be subdomen on my ciffirder may be renonically brings bone but on this, expec
7. Sherlock Holmes and the ussign and the chair of who, did them don't yourse foll!" "A consider was murt if Geery rather than
8. Sherlock Holmes and the high brotherw, legating thropilase him. "I am uberead to have all come our claimin's reviouselmans,
9. Sherlock Holmes and the Boad stunding it in a originary of pace. Course freet himself, and that it hand been matter in the h
10. Sherlock Holmes and the most colour wound by poligis; the Niew Stod pat had been uartly defending that for pa. The use also

generated with beam

1. Sherlock Holmes and the muscles of those which has been considered in their constitutions, and that they could not understand
2. Sherlock Holmes and the course of their constitution of those which they would not have been considered that there was nothin
3. Sherlock Holmes and the muscles of national constitution, and that there was not thinking that there was nothing that there w
4. Sherlock Holmes and the course of their constitution of this constitutions, and that there was not understanding through the
5. Sherlock Holmes and the muscles which they could not constitution, and they were that there is nothing that that there is not
6. Sherlock Holmes and the course of their constitution, and there is nothing that they would have been considered that they had
7. Sherlock Holmes and the pressure of their constitution of those which has been constitutionally. There is nothing that there
8. Sherlock Holmes and the country of their constitutionation of the country, and that there was not followed by those of their
9. Sherlock Holmes and the constitutions of their constitutions, and that they have been followed that there was nothing by the
10. Sherlock Holmes and the course of those which there is usually considerable, and that they could not know that there was not