

# LING 575K HW8

Due 11:59PM on May 23, 2024

In this assignment, you will

- Develop an intuition of the flow of recurrent Seq2Seq models
- Implement key pieces of the model architecture
- Better understand the formal process of attention
- Learn and practice writing useful Python docstrings

All files referenced herein may be found in `/dropbox/23-24/574/hw8/` on patas.

## 1 Understanding Attention [20 pts]

### Q1: Padding and Softmax

[3 pts]

- What is  $e^{-\infty}$ ? Or more accurately, what is  $\lim_{x \rightarrow -\infty} e^x$ ?

**Answer:**

$$e^{-\infty} = 0$$

- Give the equation for softmax over an input vector  $x$  for the value  $x_i$

**Answer:**

$$\text{softmax}(x_i) = \frac{e^{x_i}}{\sum_{j=1}^n e^{x_j}}$$

- What would happen to the softmax output if every  $x_i$  in  $x$  were set to  $-\infty$ ?

**Answer:**

Given our answer for the first part of this question, we can surmise that the expression  $\sum_{j=1}^n e^{x_j} = 0$  if every value of  $x = -\infty$ . Therefore, this would lead to a situation where we would be dividing by 0 in our softmax function as defined above. This would mean that we would get an undefined result.

**Q2: Attention Calculation** Complete the computation of attention given the following values for Queries and Keys. Interpret the Queries and Keys as having a hidden dimension of 2, with the other dimension being sequence length. Following the case of RNN Seq2Seq, Keys and Values are the same. You can use a library like Numpy or Pytorch to do the computations, but make sure to do them step by step. When filling in the matrices, round to two decimal places, but do not round in your actual calculations. [10 pts]

$$\begin{array}{cc} \begin{bmatrix} 0.2 & 0.4 \\ 0.8 & 0.6 \\ 1.0 & 1.2 \end{bmatrix} & \times \begin{array}{c} \begin{bmatrix} 0.1 & 0.7 & 0.9 & 1.5 \\ 0.3 & 0.5 & 1.1 & 1.3 \end{bmatrix} \\ \text{Keys}^T \end{array} = \begin{array}{c} \begin{bmatrix} 0.14 & 0.34 & 0.62 & 0.82 \\ 0.26 & 0.86 & 1.38 & 1.98 \\ 0.46 & 1.3 & 2.22 & 3.06 \end{bmatrix} \\ \text{Compatibility} \end{array} \end{array}$$

$$\begin{array}{c}
+ \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -\infty \\ 0 & 0 & -\infty & -\infty \end{bmatrix} \\
\text{Mask}
\end{array}
=
\begin{array}{c}
\begin{bmatrix} 0.14 & 0.34 & 0.62 & 0.82 \\ 0.26 & 0.86 & 1.38 & -\infty \\ 0.46 & 1.3 & -\infty & -\infty \end{bmatrix} \\
\text{Masked Compatibility}
\end{array}
\begin{array}{c}
\begin{array}{c} \xrightarrow{\text{softmax}} \end{array} \\
\begin{bmatrix} 0.17 & 0.21 & 0.28 & 0.34 \\ 0.17 & 0.31 & 0.52 & 0.00 \\ 0.30 & 0.70 & 0.00 & 0.00 \end{bmatrix} \\
\text{Attention}
\end{array}$$

$$\begin{array}{c}
\times \begin{bmatrix} 0.1 & 0.3 \\ 0.7 & 0.5 \\ 0.9 & 1.1 \\ 1.5 & 1.3 \end{bmatrix} \\
\text{Values}
\end{array}
=
\begin{array}{c}
\begin{bmatrix} 0.92 & 0.90 \\ 0.70 & 0.78 \\ 0.52 & 0.44 \end{bmatrix} \\
\text{Output}
\end{array}$$

### Q3: Attention Questions

[7 pts]

- Give the equation for attention (i.e., the Output of Q2) in matrix notation, using Q, K, and V to represent the Queries, Keys, and Values matrices (no need to show actual numbers).

**Answer:**

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V$$

- In the Compatibility matrix, what does each row represent?

**Answer:**

Each row in a compatibility matrix corresponds to a specific decoder time-step, with each value therein representing the importance or relevance of each encoder state for producing the current decoder state.

- Why do Keys and Values have to be the same length?

**Answer:**

The Keys and Values have to be the same length because the softmax of the product of the Keys and Queries (i.e. the Attention) needs to be able to multiply with the Values. This is only possible if the number of columns in the Attention matrix is equal to the number of rows in the Values matrix.

- In a recurrent encoder-decoder model, where do the Queries, Keys, and Values come from respectively?

**Answer:**

Queries come from the decoder's hidden states. They represent the current hidden state of the decoder at each time step of the decoding process.

Keys and Values come from the encoder's hidden states. Keys encode information from the input sequence and are used to determine the relevance of each input token to the current decoding step. Values contain the actual information from the input sequence that will be weighted and combined to form the context vector for the decoder.

## 2 Documenting Code [10 pts]

Docstrings are an important part of Python development, and so are an important part of developing a Deep Learning system. Use these links as a reference for Google style docstrings, which we will use here [https://sphinxcontrib-napoleon.readthedocs.io/en/latest/example\\_google.html](https://sphinxcontrib-napoleon.readthedocs.io/en/latest/example_google.html), <https://google.github.io/styleguide/pyguide.html#s3.8.1-comments-in-doc-strings>. You can also use the docstrings in `model.py` as examples. This format has the advantage of being able to be used to automatically generate html documentation from code for a project's website. Here, you will provide docstrings for the `Seq2SeqDataset` class in `data.py`, as well as its methods `batch_as_tensors` and `from_files`. Please only include the main description, **Args**, and **Returns**, giving roughly the level of detail in `model.py` (as opposed to the link).

The following section of the document is an environment that allows you to type monospaced plain-text in Latex. If you are using Latex, please make sure to keep your docstrings in these environments. If you are submitting a .txt file, you can just use `"""` to demarcate the beginning and end of your docstrings. This environment will wrap your lines for you, but try practicing manually formatting your lines to be 80 characters. Most text editors will allow you to show a ruler line at a specific number of characters, so try drafting your docstrings there and pasting them here.

```
class Seq2SeqDataset(Dataset):
```

```

"""
A instance of the Dataset class meant for Seq2Seq data modeling.
Arguments:
    examples: A list of Examples, defined above as a dictionary mapping strings to any
              object
    vocab: A Vocabulary instance, as imported from vocabulary.py
"""

def batch_as_tensors(self, start: int, end: int) -> dict[str, Any]:
"""
Generates a dictionary of tensors from the examples passed into the model between a start
and ending index.
Arguments:
    start: the lower index of the range of examples from which the dictionary of tensors
           will be generated
    start: the upper index of the range of examples from which the dictionary of tensors
           will be generated
Returns:
    A dictionary containing the tensors for the source, target_x, and target_y as well as
    the lengths for the source and target tensors.
"""

@classmethod
def from_files(
    cls, source_file: str, target_file: str, vocab: Vocabulary = None
):
"""
Generates an instance of Seq2SeqDataset using an input file.
Arguments:
    source_file: a file representing the source data, represented as a continuous string.
    target_file: a file representing the target data, represented as a continuous string.
    vocab: the Vocabulary object to be used by the dataset.
Returns:
    A Seq2Seq dataset using the input files as examples.
"""

```

### 3 Implementing a Recurrent Seq2Seq Model [37 pts]

See hw8.tar.gz

### 4 Running the Translation Model [8 pts]

`run.py` contains a training loop to use the Seq2Seq translation model, which will record the training loss and generate text every  $N$  epochs (controlled by `--generate_every`, set to 1 by default). **Warning: the run of this model will take several hours or more. It is important that you plan ahead, and only use the default parameters of the script (which are small), until you are sure your code runs end to end without errors and it is time to run your final script.**

**Q1: Run with Full Parameters** Execute `run.py` with the following (full) arguments. Paste below the texts that are generated every epoch, as well as the training and validation loss. In 2-3 sentences, describe any trends that you see. [4 pts]

- `--train_source /dropbox/23-24/574/data/europarl-v7-es-en/train.en.txt`
- `--train_target /dropbox/23-24/574/data/europarl-v7-es-en/train.es.txt`

- `--output_file test.en.txt.es`
- `--num_epochs 8`
- `--embedding_dim 16`
- `--hidden_dim 64`
- `--num_layers 2`

### Answer:

The generated outputs in the first few epochs, and in the first epoch especially, feature very small word fragments with hardly any coherence to them. As the epochs proceed, we start to see words and phrases increase in complexity. Though it seems do get a bit caught in recursion (especially in the “de la comisión” sections) it begins to at least make coherent Spanish sentence fragments by the last epoch.

### Output:

Epoch 0 train loss: 2.8618

Epoch 0 dev loss: 2.3338; perplexity (nats): 10.3175

New best loss; saving current model

Generating example output:

INPUT: is this not, indeed, what is so distressing to some socialist members of the european parliament?

GENERATED OUTPUT: la le le de de tanta de er en antes en le ante de puen en pue de de de les de pue de da da puen en en en en ente de de de en la tes de pocon en ente con pue pues es en en nente en en le les en en en es lo de de pen de puen mes los enta en en de de la de de entema de en de de de de de ente ente de peste de pon en de de le ten de pon en en ente la de enta de de con le de do en le ente per da en en la da pon en ente da de des ente cen en en le les de de en al quo en la de le teres de la de puen de pes puen le

INPUT: mr president of the commission, our objectives are one and the same: we want a strong europe in the twenty-first century, a democratic europe which is able to act.

GENERATED OUTPUT: ra en en en res de puo de an en en de les en le en cuen en al de le des en ente to de de en la curo de le en er de en en en les en en ecte en ente de de puen en en en en en ente en cos en de en pes ente es es en en en le de le cure de de la de de da de de da puente por en en ente de de de de cor de con la de le le de da de le de le de cos en en le de de pores con en de de cos de cue con en de lo de toc en en en en es en suen en sos de sen la pue le do len en sosta en es entos de le con en en los es pen

INPUT: we must take up a reasoned offensive against radicals on both the right and the left.

GENERATED OUTPUT: es pue le ente sos de poren en en de le en en ente do en en de la de en en en le la en en lo puue an enta le de de cante en er ente pon les puen es en en enta en en en la de la pen pen es cuen er per conen la en en de la peran en el de de de que por de pue en en las de cos er la de de de cura cien en de pren en en los de les en de les en la les de pue de de ante de es en puen la ente conta les de de le de tara de pues les en en en le des le la en en en en en es puen en en es ecuen es cos de la le puen en

INPUT: in this respect, one of the variants considered is reminiscent of the dutch polder model, the structurally planned cooperation between government, trade unions and employers' organisations.

GENERATED OUTPUT: es de por en ento de en en la de la puente en pero en cure en en pe de pos la de de en en en les sos de de en en en en en ante en de en le de en en en en er cen es en ce cier de en en en con en en ente a en en a le la tos code en en en en con de en en en en en lema en en pues pres en e les le ante en en en la en que en de de de en son an sos de con den en en en en les en es de en la de de de en la de le por puente pon ente

le en ente ces es pue con en le les de to de de la ante de la puen le

INPUT: sometimes, an enormous reservoir of money is created, not because it is not desperately needed, not because there are no sound proposals, but because the whole system has collapsed.

GENERATED OUTPUT: e re sune ente de por de en de de la de de le pen en da de con en la pon en la de per de de pen ses en le de en ente te to de con la en sura en la de en ente en de la en conte de en pue en an en en qte de de de le conte con es en la de per la les anto le der con le de pue de de an de de la das de de con les en en aner en de sos de en esta en en er de cue de en en en por ente an en de en en les men ente en ente en entes puen en entar pos de le ante cen en en la en le la de an le ses en en les en en en le de

INPUT: i do not have these concerns.

GENERATED OUTPUT: es en en es ente de des en en pue yes pue de de en sen en en de en a den en en suen cos de de les cades de pon ente es de de cacon en peren en an es les ente en en desa en en de la de les ente cuen en en en en de puen en en de de an les en en puen la da en es en en en cecos en la de pen es de la de des prestar en en sos de puen de ente pues pera ante de de de en en en en en en en en ser en en en enes que pue de de de en el de en en en sures en la pos de de de el de cena de pante de de de por ente cien

INPUT: secondly, regulation 1762 must be repealed and replaced with a new regulation.

GENERATED OUTPUT: es en en en en en de en cura de de da de le cun en en en con en en de esien ente les en en entes de en de puen en en en en en ente pen en en pues en ente de de la de cin ente en sos de en es en de en en en en de per es la la la en en puen en en mes la de an la de des en en en de pue pien de en en en es en la de en en es pues er de pues conte de la de curoc da de puede de des en en la de puen en en an en en en en en ente pue le con en de de ponte de pures de cucien en en en en lo de les puen en en en la pred

INPUT: the interreg initiative must include eligible areas with special emphasis on border areas, islands and both mountain and non-mountain areas, such as the prefecture of arta in greece, whose omission from annex i of the commission's communication is ill-considered.

GENERATED OUTPUT: e la de en en la de dos es cura en cos enta cos la so puen en cuse la de de sente de en puen de pon el pon en en le puen en re conten es so de de puen en en en de cos en en ente la de so de de de da pue ente cor de pon car de con ente es en en de le puen en en este en en en la de en en de de pen en en ente pe en en pue cua de en de de la de ente da te de pres de en en ente en cua en la en en er sos en en ente mo cue de pon en ente en perien la le de de la con la de nente en emen en en en ente en nes en sor

Epoch 1 train loss: 2.1911

Epoch 1 dev loss: 1.99; perplexity (nats): 7.3152

New best loss; saving current model

Generating example output:

INPUT: is this not, indeed, what is so distressing to some socialist members of the european parliament?

GENERATED OUTPUT: el en esta de en la en en astamente en la preses de la en este de la de esto de la contamico de la en la de la la la acures de la en en en en la que en la de la entrente de la preses de la de la en esta de prención de la en en en en en esta contrente de la condes de sesinte de los en en en los de la contento en la cuenta de la de la que la la en en la contente de la conteres de la conconte en en en en la en la conserta en en en la la seminte de la preses conteres de la desente en la preses de la sustente es

INPUT: mr president of the commission, our objectives are one and the same: we want a strong europe in the twenty-first century, a democratic europe which is able to act.

GENERATED OUTPUT: en este en la presere de la la de esta la de la se en acuras en la ciunes en en conteridas de la conteres de sesta de el precidal de la en este de la premente en en la pares de la la camos de la de la en pregición de la la presidente en se en se los presinte

en la precisión de en la contenta de un presición de los de de la tomenta de la conterente este de la contento el la de la la de en la en la sesta de la contiminta de la sen en el presición de la la de un en este de la la en en en sesición de la de la d

INPUT: we must take up a reasoned offensive against radicals on both the right and the left.

GENERATED OUTPUT: se de la de conteres de en contara en la desterta de la en en prente de la de la de la en en en la en en presición de la de la de el properes este concición en en en en la de la conterente es en presidenta en la presición de la la en en en la de la sinte por este esto en este cuento de en la precianta en en el contenido de tente de en la puesta de el de la mente y en el la que en este de la la de la en esto en la promente esta de la de en este de el precipar esta en en en un presición de en en en en en en pr

INPUT: in this respect, one of the variants considered is reminiscent of the dutch polder model, the structurally planned cooperation between government, trade unions and employers' organisations.

GENERATED OUTPUT: se el compresición de la sección de este de la presición de la preción de la la de la en en la prades de prente en la de terente de la en el premente de la peres la en la entres de la soma de la la sestación se este de la de la de la de la de la conteres de la la de la prente de la presición de la la en en en en en la acuranta en en es en el prenta en en el presiria en en en la contera en la la sestras de la prente de la en la en de la senta en la conteres con de la de la contenta de la en en en la en este e

INPUT: sometimes, an enormous reservoir of money is created, not because it is not desperately needed, not because there are no sound proposals, but because the whole system has collapsed.

GENERATED OUTPUT: en la de la la de la esta de la en la de prostados en la contenta en estima es estar en la la en esta de la en en la en contenido de la la contrente de la de en en este de en la la en se tares de la la de la en en la mente de la en se la la contonta en en la sesto en la contorde en este en la presente en la la en la en en se se de seminas de en en sesides de la en esta de la la esta de el presición de la de la la conteridas es este de la conteres de la esta en la de la la de la en la en en en la la de la ses

INPUT: i do not have these concerns.

GENERATED OUTPUT: se la de la la contente en en la entera en en en este de la preres en entos en prente de la secianido a la de este de la conteres de en la la contidente en la la este de la la la presición de la estas de la la de la de astenta conteras de la conteria de la cominida comento en sesta de la enteridas de la en en la mobre por en en la la en la de la de la las de la seminto en en la en en la que la presición de la estas en en la la cuente en en este de el prenpos en sostar a la en el preseridas de la la en esto

INPUT: secondly, regulation 1762 must be repealed and replaced with a new regulation.

GENERATED OUTPUT: esta de la la en en la la la entere con esta en conteres de la de contenido de la semos de la cuente de la la de este de la presición de la que en en la de en la la en de la prente de en en en cuente en la en en la la de la de la con de en en este en la estar este de sestida de la concuante son en la la de la de la prentadas se la en esta en esta de la este en la la sestar esta conteres de la contenido de en la de la acuenta el preci3n de la en en la propires en en la cinto de en la de la de en la en entera e

INPUT: the interreg initiative must include eligible areas with special emphasis on border areas, islands and both mountain and non-mountain areas, such as the prefecture of arta in greece, whose omission from annex i of the commission' s communication is ill-considered.

GENERATED OUTPUT: es esta de la la la de los esta de la de en en el preستا de la que en en en resente de la la la la conteres de la en en en la contera en la en en la la de la contide en el de la conterta en se contar en en un a la contenido de la en el la cuente de en la en el de en en contres de la eminta de sestar de la en la la cación de la de cuente de la la contenta de la la desen en este en la de la la contenido es en en la contenido en la senta en la de sus de en se esta de la pobre en la en acura con la en tortera d

Epoch 2 train loss: 1.9864

Epoch 2 dev loss: 1.8495; perplexity (nats): 6.3566

New best loss; saving current model

Generating example output:

INPUT: is this not, indeed, what is so distressing to some socialist members of the european parliament?

GENERATED OUTPUT: esta presición en esta de la precición de la comini3n a la unesi3n el presidente de la comines que estar el presimidad de la presiciones en esta prantes de la presidente de la presición de la presición de la precirar esta de la presición de la presición de la comini3n de la presición de la presición de la presición de la presición a la comente de la decición de la precidente de la contresión de la concuento de la presiciones de la presenta esto preستا de en esta presidera de la presición de la del preside

INPUT: mr president of the commission, our objectives are one and the same: we want a strong europe in the twenty-first century, a democratic europe which is able to act.

GENERATED OUTPUT: en de por la desenti3n de los presiciones de la presición de la presididad de la parti3n de las camini3n de la presidente de la de presidente en el presición de la partes de la presición de la parti3n en acuera para este contros presidera el contraci3n de la contresión de la presiti3n de la presición de la presición de la paresi3n de la presición en un precidente de la presición de la presidente de la presidente de los presinicadados en esta presición de la presiti3n de la presición a la presición de la

INPUT: we must take up a reasoned offensive against radicals on both the right and the left.

GENERATED OUTPUT: esto de la presides de la presición presidente de la preciones de la presición este presentado este presiti3n de la presición de la de la presición de la presición de la presición con la del presición de la presición de la presición de una presición de la desimidad por el presidente en en la presiciones de estamos de la compera en esta de en presición de el presimica de la presición de la presición esta presición de la presición de la presicia este parte de la comisi3n de la presición de la tortici3n de la

INPUT: in this respect, one of the variants considered is reminiscent of the dutch polder model, the structurally planned cooperation between government, trade unions and employers' organisations.

GENERATED OUTPUT: se presición de la presición de la contertass de la presitiende de la presición de la presidente de la presimidad de la presidente de la semici3n de la presidente de la presicion estado en en esta presiciones de la presenta de la entresión de la presidente de la presidente de la contresión de la presición de la presición de la presimida se por la contresión de la presición de la presición de la presiciones de la presición de la presición de la presición de la presición de la partes de la presimica de la co

INPUT: sometimes, an enormous reservoir of money is created, not because it is not desperately needed, not because there are no sound proposals, but because the whole system has collapsed.

GENERATED OUTPUT: se en esta de la portente en esto presentar esta de de la parte de la dentero en este prodes a el en la propiente de acurar en el presición de la comini3n de la presición de la superes por presidente de la pertes por en la presición de la presición en el preciente de la propuesto de la comente de la compar en la presidición en el presición de la presaci3n de la presiciones en esto de la contresión de la partente en la presidente de la entrama en esta de la por esta de la secciones de la contares en el presic

INPUT: i do not have these concerns.

GENERATED OUTPUT: esto presición de la preciania en el presidicia de la presidente de la presición de la concuento de la presición de la preciciante de la presición de la presición de la presición para presición de la contridadar el por presidente de la compente de la presición de la predidicar de la presición de la contresión de los presición de la presición en este presición de la



propuesta de la contresión en esta contramente de la preciante de la entrestar el presiciones en en la contresión de la presición de la presició

INPUT: secondly, regulation 1762 must be repealed and replaced with a new regulation.

GENERATED OUTPUT: esto par esta presición de la de una propuesta en en la contara en el presición de la presición de la compuesta en en esta de la presinte por la presición de la presitió de la presición de la presición de la presición de la presición de la presición de la comisión de los precedente con la presición de la presición de la presitió de la presición de la el presente de la presición de la compresición de la presición de la cominió de la comisión de la presición de la presición en el presición de la presició

INPUT: the interreg initiative must include eligible areas with special emphasis on border areas, islands and both mountain and non-mountain areas, such as the prefecture of arta in greece, whose omission from annex i of the commission' s communication is ill-considered.

GENERATED OUTPUT: esta menta de apresido en esta presidente de la entramente en el presición de la presición de la premente de la presición de la presición de la presición de la presición en esto de la precición con la presidente es en en en este esta acuenta por la contara en esta presimas de en en en esta presiciones de la presición de la presitió de la comperes de la presición de la presición se en presición de la la contresión de la presición de el presición de los pares presición de la propuesto el puesta de el parte d

Epoch 3 train loss: 1.8835

Epoch 3 dev loss: 1.7605; perplexity (nats): 5.8154

New best loss; saving current model

Generating example output:

INPUT: is this not, indeed, what is so distressing to some socialist members of the european parliament?

GENERATED OUTPUT: se presicia en esta este de la presición este esta de la entros presiciones de la comisión de la consentar a la entre acunte de la comunicada este de los mecto de la considente de la comisión de la comisión de la desiciones esta de esta de se comisión de la presición de la contra entros paresta de la contranter de la comisión de la del presición de el presición de los presición de la comperiones de la decienta en el para la comisión de la comisión de la comisión de la presición con la cantidad el presición

INPUT: mr president of the commission, our objectives are one and the same: we want a strong europe in the twenty-first century, a democratic europe which is able to act.

GENERATED OUTPUT: esta de la comisión de la presición de la contra esta por la comisión de la desento de la presición de la presidente, de la contras en el contra esta de la contras la encontante en el comisión se se presicica de la las parte de la deciante en esta de la comisión en la contra en esta de la considición de la dena de la contros de la comisión de la presición de la considente de la contradica de la consición de la presines al de la comisión de la presiciones en esta de la consiento de la presiciones de la del p

INPUT: we must take up a reasoned offensive against radicals on both the right and the left.

GENERATED OUTPUT: esto de la unión de la contra a la consición de la contras de en esta de la entrantica de la comisión de la comisión de la presición de la contras de la comisión de la presición de la decienta con la comisión de la comisión de la consectar la consición de esto propuesto de la presición de la tresición de la consición en en esta de la contranter a la presición de la estamos en esta de la presición en esta presición en el presición de la comisión de la presición de la estamente de la comisión de la presición

INPUT: in this respect, one of the variants considered is reminiscent of the dutch polder model, the structurally planned cooperation between government, trade unions and employers' organisations.

GENERATED OUTPUT: esta de la comisión de la comisión de la interenciones de la presimenta

de la consición de la sentar en en el propuesta el consericia en esta de la presiciones de la tramos en en esta en el consisión de la comisión de la del presitica a la presición de la contresión de la comisión de la comisión de la presición con este contresión de la de la consientas de la parte en en esta en esta de la considente de la comisión en esto presente de la consición de la comisión de la presiciante de la campera de la pramica

INPUT: sometimes, an enormous reservoir of money is created, not because it is not desperately needed, not because there are no sound proposals, but because the whole system has collapsed.

GENERATED OUTPUT: es esta esto de la comisión por la presición de la estatidad de la consición de la conciadente el pristente en la comisión de la trosipara con el consecta de la interterado de la presición esto presición de la contrados en el presidente de la comisión de la consiciones de la comisión de las contras de la estas de la denemente de la consiciones por la presimico de la consición de la consuctones de la contras este esta sin en el presición de la entres de la contresión de la estación se presidente el presición

INPUT: i do not have these concerns.

GENERATED OUTPUT: esta comisión de la presición de la comisión de la comisión de la presición de la comisión de la presiciar a la presición y la decuresión en el propuesto de la comisión de la comisión de la consición de la considente de la comisión de la paresicia de la estamente de los presición en esta estabo de la comisión de la cominica de la contros en la comisión de la prosición en esta consiciones en este presento que contra acciones de la comisión de la contración de la decianciones en esto esta de la hemos consterio

INPUT: secondly, regulation 1762 must be repealed and replaced with a new regulation.

GENERATED OUTPUT: es esta de la cambiión de la consición de la presición de la presición de la comisión de la presición de la comisión de la consición de la comisión de la contros de la comisión de la heciones de la comisión de la furando de la comisión con esta de la consiciones de la comisión de la presición de la comisión de la presición de la comisión de la comisión de la esta de el comisión de la comisión al presición de la presición de la unión en el consinidad de la comisión de la presición de las comisión de la presic

INPUT: the interreg initiative must include eligible areas with special emphasis on border areas, islands and both mountain and non-mountain areas, such as the prefecture of arta in greece, whose omission from annex i of the commission' s communication is ill-considered.

GENERATED OUTPUT: es esto consiciones de la contares en esto presiciones de la comisión de la entra esta la presición de la esta de la comisión de la presición de la consición de la interesión en esta de la consición de la presición de la consisión de la consectado en el consepiales de la entras a la interteración de la precente de la presiciaciones en la comisión de la consición de la comisión se para la estación de la consición de la contra en la considicia de la comisión de la entros prantica en la presición de la comisi

Epoch 4 train loss: 1.8155

Epoch 4 dev loss: 1.7031; perplexity (nats): 5.4909

New best loss; saving current model

Generating example output:

INPUT: is this not, indeed, what is so distressing to some socialist members of the european parliament?

GENERATED OUTPUT: esta de la comisión de la comisión de la contros de la comisión de la presicial de la propuesto de la contras en priciones de la comisión de la consición de la presento de la propuesta de la consición para la comisión de la comisión de la contras de la presidente de la contra la consición a la comisión de presiciones de la comisión de la comisión de la presidente de la presición de la comisión de la comisión en el properante de la comisión de la comisión de la presicidad de la comisión de la comisión en est

INPUT: mr president of the commission, our objectives are one and the same: we want a strong

europe in the twenty-first century, a democratic europe which is able to act.

GENERATED OUTPUT: este precente de la comisión de la considente de la comisión de la presento de la presidente en la comisión de la propuesto de la comisión de la pranter en el probre para la comisión de la comisión de la comisión de la propuesta de la presimiva de la contra de consecciones de la presidente de la presidente de la contras de la presento de la comisión de la comisión de la consición de la consección de la consición de la comisión de la comisión de la comisión de la comisión de la comisión de la presición de la

INPUT: we must take up a reasoned offensive against radicals on both the right and the left.

GENERATED OUTPUT: esta contro presición de la consición en el propuesto de la comisión de la comisión de la contra de la consición de la presición de la comisión de la comisión de la contras de la comisión de la cambiión de la considera de la comisión de la comisión de la comisión de la comisión de la propuesta de los presiciones de la comisión de la comisión de la presición de la comisión de la comisión de la presición para la contros de la consición de la propuesta de la comisión de la comisión de la comisión de la comisión

INPUT: in this respect, one of the variants considered is reminiscent of the dutch polder model, the structurally planned cooperation between government, trade unions and employers' organisations.

GENERATED OUTPUT: señor presiciones de la comisión de la presición de la comisión de la comisión de la comisión el consertan en la considera de la de la comisión de la comisión de la comisión de la comisión de la comisión de la comisión de su presicidad de la presiculicad de la comisión de la comisión de la comisión de la comisión de la presicidad de la comisión de la considente de la comisión de la presición de la comisión de la considente de la comisión de la del presicicia de la comisión de la comisión de la comisión de l

INPUT: sometimes, an enormous reservoir of money is created, not because it is not desperately needed, not because there are no sound proposals, but because the whole system has collapsed.

GENERATED OUTPUT: esta comisión de la propuesto de la comisión de la comisión de la contras de la trosidad de la comisión de la contra de la propuesto de la presición de la comisión de la contras a la comisión de la comisión de la comisión de la comisión de la contra la presicion de la comisión de la contra entra consiciones de la comisión de la consición de la comisión de la presidente de la comisión de la comisión de la contra a la comisión de la contra contra el presidente de la comisión de la comisión de la comisión de

INPUT: i do not have these concerns.

GENERATED OUTPUT: señor presiciones presiciones de la considente de la propuesta por la presición de la considente de la considente de la comisión de la comisión de la consición de la consuciones de la comisión de la propuesto de la comisión de la comisión de la comisión de la consicial de la comisión de la comisión de la considente de la presición de la comisión en la presición de la propuesta de consertar a la comisión de la comisión de la comisión de la presiciones de la comisión de la contrasitados de la consición de la

INPUT: secondly, regulation 1762 must be repealed and replaced with a new regulation.

GENERATED OUTPUT: es el presición en la contra del contros de la comisión de la comisión de la consición de la comisión en considera de la constición de la comisión de la propuesto de la comisión de presimidad de la comisión de la comisión por la comisión de la comisión de la comisión política de la presición de la presición de la contros de la contrante de la consición de la compresente con la comisión de la comisión de la propuesta en en el propuesto de la presimidad por la presidente de la comisión de la presición de la c

INPUT: the interreg initiative must include eligible areas with special emphasis on border areas, islands and both mountain and non-mountain areas, such as the prefecture of arta in greece, whose omission from annex i of the commission' s communication is ill-considered.

GENERATED OUTPUT: en esta presenta en la comisión de la considente de la comisión de la comisión de los presiciones de la comisión de el miento de la comisión de la consiciones de la propuesta

de la presidente de la comisión de la propuesta de la consición de la comisión de la comisión de la propuesta de la comisión de la comisión de la comisión de la considente de la comisión de la propuesta de la presición de la comisión para la comisión de la comisión de la comisión de la comisión de la presicidad de la considente de la c

Epoch 5 train loss: 1.7654

Epoch 5 dev loss: 1.6525; perplexity (nats): 5.2203

New best loss; saving current model

Generating example output:

INPUT: is this not, indeed, what is so distressing to some socialist members of the european parliament?

GENERATED OUTPUT: esta presiculidad de la comisión de la comisión en el presento de la comisión en el contra de la comisión de la comisión en el propuesta de la comisión de la comisión de la comisión de la presicia de la consienta de la comisión de la comisión de la comisión de la comisión de la comisión de la comisión de la comisión para la comisión de la comisión de la comisión de la comisión y el presidente en en la comisión de la comisión de la presicia en la presicudidad de la comisión de la comisión en la comisión de l

INPUT: mr president of the commission, our objectives are one and the same: we want a strong europe in the twenty-first century, a democratic europe which is able to act.

GENERATED OUTPUT: se esta presentar en el presidente en el presento para la comisión de la comisión de la consicia de la comisión de la comisión de los presidente en el presenta de la comisión de la presicia en el contra de la comisión de la presidente de los presiculicadas de la comisión de la presidente en el presidente de la comisión de la presidente de la comisión de la comisión de la comisión de la comisión de la comisión de la presiculicados de la presicia de la presicia en el presidente a la comisión de la comisión de

INPUT: we must take up a reasoned offensive against radicals on both the right and the left.

GENERATED OUTPUT: la dema de la presiciones de la comisión de la contras de la comisión de la comisión de la comisión de la comisión de la comisión en esta presicia de la comisión de acuerda de la comisión de la presidente de la comisión de la presidenta en la presiciones de la consinicado en el presiciones de la comisión de la comisión de la comisión de la comisión de la comisión de la comisión de la comisión de la conseración de la propuesta de la comisión de la comisión de la comisión de la comisión el presiciones de la c

INPUT: in this respect, one of the variants considered is reminiscent of the dutch polder model, the structurally planned cooperation between government, trade unions and employers' organisations.

GENERATED OUTPUT: señor presicia en el contra consiciones de la contras de la propuesta de la presición de la comisión de la comisión de la comisión de la presición de la propuesta de la comisión de la comisión de la comunidad de la comisión de la comisión de la presiciones con la comisión de la comisión de la comisión en la comisión de los presiciado en presiciones de la comisión de la comisión de la comisión de la comisión de la comisión de la contras en el presiciones de la comisión de la comisión de la comisión de la com

INPUT: sometimes, an enormous reservoir of money is created, not because it is not desperately needed, not because there are no sound proposals, but because the whole system has collapsed.

GENERATED OUTPUT: este presicicia de la interesión de la comisión de la comisión de la comisión de la comisión de la comisión de la comisión de la comisión de la comisión con la comisión de la comisión de la consideres de la comisión de la comisión de la comisión de la comisión de la comisión de la comisión de la consicuales en la presicia para la comisión de la procesto por la comisión de la contras de la comisión de la comisión de la comisión de el propuesto para la comisión de la contración de la comisión de la comisión d

INPUT: i do not have these concerns.







[illegible]

**Q2: Evaluate Translations** Finally, we will evaluate the translations using Character F-Score (<https://www.aclweb.org/anthology/W15-3049/>). Using the provided script `chrF++.py`, calculate the score of the output Spanish translations using the following command, and report the score for the line that says `c6+w0-F2`: [4 pts]

```
python chrF++.py -nw 0 \
-R /dropbox/23-24/574/data/europarl-v7-es-en/test.es.txt \
-H test.en.txt.es > test.en.txt.es.score
```

**F-Code:**

c6+w0-F2 15.5228

## 5 Testing your code

In the dropbox folder for this assignment, you will find a file `test_all.py` with a few very simple unit tests for the methods that you need to implement. You can verify that your code passes the tests by running `pytest` from your code's directory, with the course's conda environment activated.

## Submission Instructions

In your submission, include the following:

- `readme.(txt|pdf)` that includes your answers to §1, §2, and §4.
- `hw8.tar.gz` containing:
  - `run_hw8.sh`. This should contain the code for activating the conda environment and your run commands for §4 above. You can use `run_hwX.sh` from any previous assignment as a template.
  - `model.py`
  - `test.en.txt.es`