

Python 程序设计与数据科学导论（期末大作业 Part 2）

机器翻译

# 目标

1. 利用PyTorch 复现 Seq2Seq with Attention，以实现简单的机器翻译任务。

# 主要任务

1. 数据预处理 （已由助教在Dataloader类中实现）
2. 模型：Encoder、Decoder（+ Attention）、Seq2Seq
3. 测试效率（生成文本等） + 调参

目录

[模型 3](#_Toc76651616)

[1. Encoder 3](#_Toc76651617)

[2. Decoder（带Attention机制） 3](#_Toc76651618)

[3. Seq2Seq模型 4](#_Toc76651619)

[模型效果： 4](#_Toc76651620)

[参考文献： 5](#_Toc76651621)

|  |  |
| --- | --- |
| 姓 名： | 尤俊浩 |
| 学 号： | 1900094810 |
| 院 系： | 信息科学技术学院 |
| 联系方式： | 1900094810@pku.edu.cn |

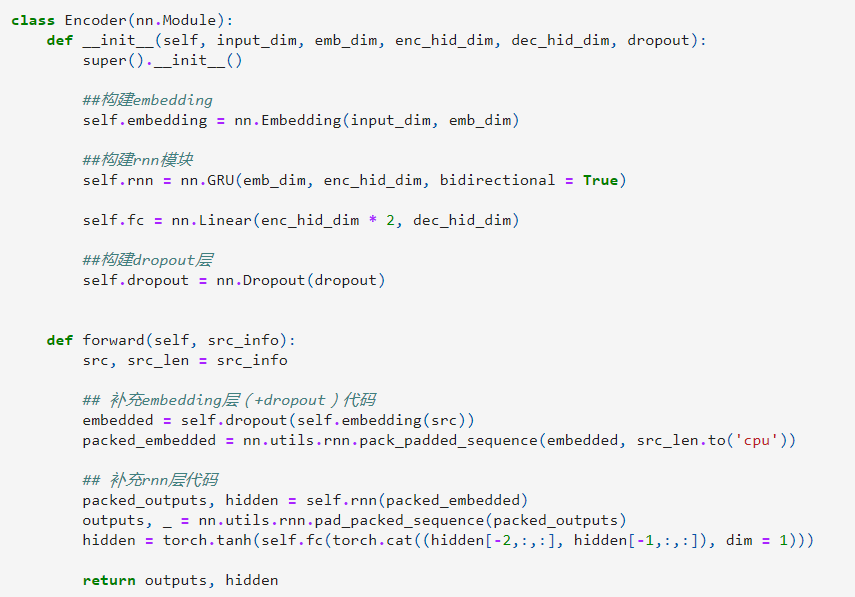
二〇二一年 六月

# 模型

模型主要有3个部分

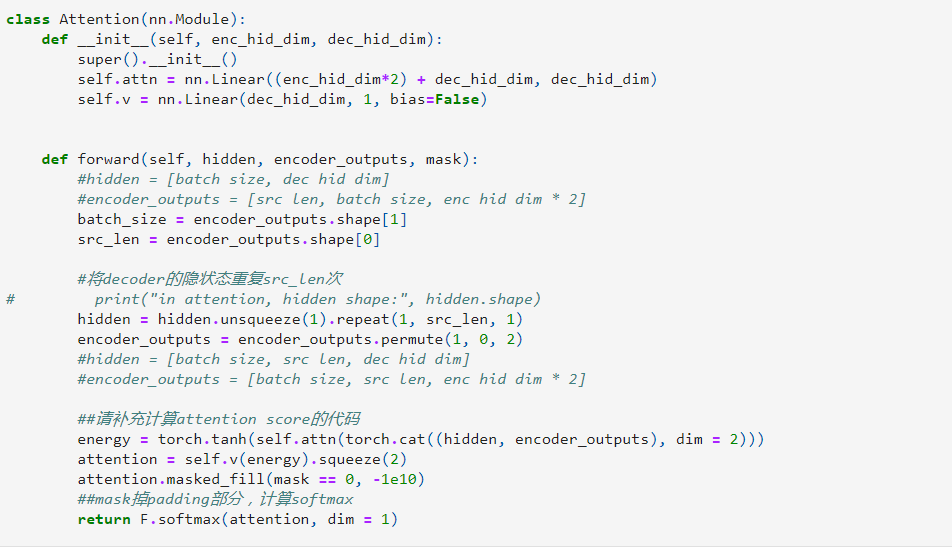
## Encoder

采用单层双向GRU，需要注意的是要在最后通过一个全连接层将输出的维度转换成[batch\_size, dec\_hid\_dim]



## Decoder（带Attention机制）

Decoder与Encoder类似，采用的是单层单向的GRU，但增加了Attention机制，以实现对序列上下文的重点进行关注。Attention机制的实现如下图，主要思想是按照元素重要程度加权平均。

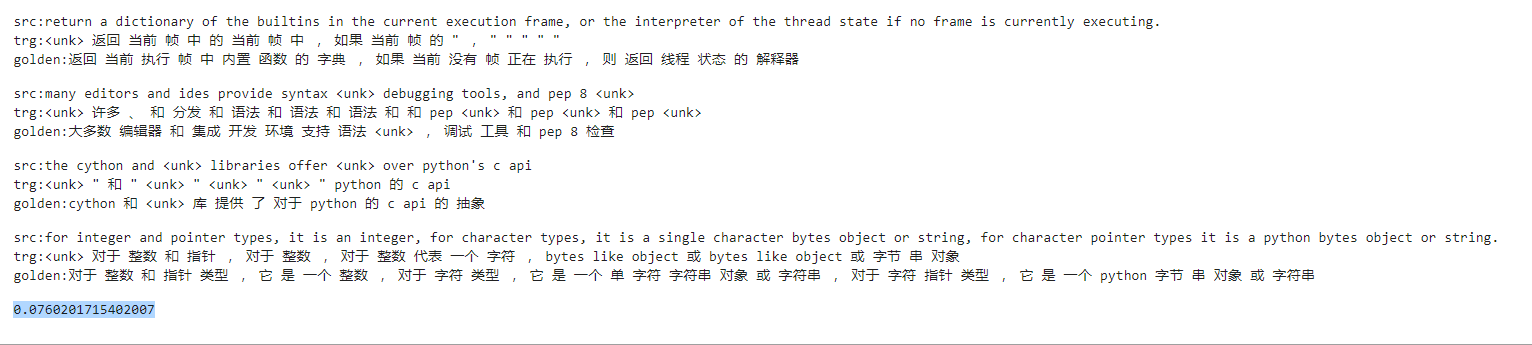


## Seq2Seq模型

最后是整合了Encoder和Decoder的Seq2Seq模型。有别于传统的Seq2Seq模型中直接将句子中的每个词都不断传入Decoder中训练，引入了Attention机制可以自己控制将词一个一个的输入。具体实现如下：

# 模型效果：

对训练集训练100个epoch后，表现如下：

利用测试集对最佳模型进行评估，平均BLEU score为0.076，其中翻译效果仍然差强人意。

# 参考文献：

[1] Seq2Seq (Attention) 的 PyTorch 实现，https://wmathor.com/index.php/archives/1451/，2020-07-02

[2] Pytorch-seq2seq，https://github.com/bentrevett/pytorch-seq2seq，2020-05-12

[3] Seq2Seq实现神经机器翻译，https://zhuanlan.zhihu.com/p/73141235，2020-11-07