

**深 度 学 习 实 验**

**实 验 报 告**

**实验名称：循环神经网络**

**姓 名：李岚琦**

**学 号：2111078**

**专 业：智能科学与技术**

**人工智能学院**

**2023年12月**

目录

[一、 问题简述 2](#_Toc152693267)

[二、 实验目的 2](#_Toc152693268)

[三、 编译环境 2](#_Toc152693269)

[四、 基础代码 2](#_Toc152693270)

[1. 导入所需的包和数据集： 2](#_Toc152693271)

[2. 设置训练超参数： 2](#_Toc152693272)

[3. 实例化LSTM： 2](#_Toc152693273)

[4. 训练并可视化： 3](#_Toc152693274)

[5. 根据提示词来预测： 3](#_Toc152693275)

[五、 基础要求与附加题 3](#_Toc152693276)

[1. 基础要求 3](#_Toc152693277)

[要求1 4](#_Toc152693278)

[要求2 4](#_Toc152693279)

[要求3 4](#_Toc152693280)

[要求4 4](#_Toc152693281)

[要求5 4](#_Toc152693282)

[要求6 4](#_Toc152693283)

[要求7 4](#_Toc152693284)

[要求8 5](#_Toc152693285)

[要求9 5](#_Toc152693286)

[要求10 5](#_Toc152693287)

[2. 附加题 5](#_Toc152693288)

[要求1 5](#_Toc152693289)

# 问**题简**述

利用以LSTM为代表的卷积神经网络，对time\_machine数据集进行训练，并且对于所要求的提示词进行预测。同时调整超参数，使得预测结果更合理。

# 实验目的

1. 学习循环神经网络的原理

2. 动手实现循环网络，并完成训练和预测

3. 改变超参数，观察对训练和预测结果的影响

# 编译环境

语言：python 版本： 3.7.16

编译器： vscode 版本：2023

深度学习框架 ： pytorch 版本： 1.10.1

工具包 ： d2l

# 基础代码

## 导入所需**的包和数据集**：

**import** torch

**from** torch **import** nn

**from** d2l **import** torch **as** d2l

batch\_size**,** num\_steps **=** 32**,** 35

train\_iter**,** vocab **=** d2l**.**load\_data\_time\_machine(batch\_size**,** num\_steps)

## 设置训练超参数：

vocab\_size**,** num\_hiddens**,** device **=** len(vocab)**,** 256**,** d2l**.**try\_gpu()

num\_epochs**,** lr **=** 500**,** 1

## 实例化LSTM：

num\_inputs **=** vocab\_size

lstm\_layer **=** nn**.**LSTM(num\_inputs**,** num\_hiddens)

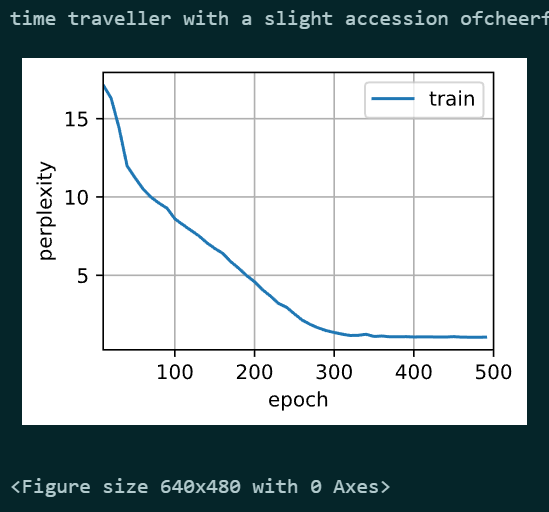
model **=** d2l**.**RNNModel(lstm\_layer**,** len(vocab))

## 训练并可视化：

      model **=** model**.**to(device)

d2l**.**train\_ch8(model**,** train\_iter**,** vocab**,** lr**,** num\_epochs**,** device)

训练的可视化过程为：



可见，随着迭代次数增加，困惑度逐渐下降，并且在epoch = 300 左右时，接近最小值。

## 根据提示词来预测：

d2l**.**predict\_ch8('time traveller'**,** 50**,** model**,** vocab**,** device)

预测结果为:



与实验示例中的预测保持一致。

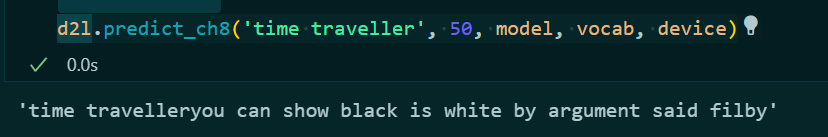
# 基础要求与附加题

## 1. 基础要求

对于给定的提示词，完成模型的输出。

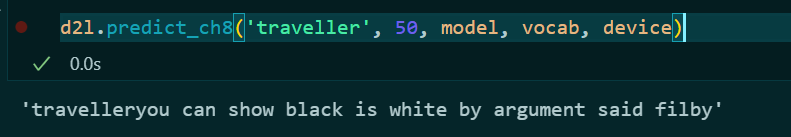
### 要求1

提示词为：time traveller



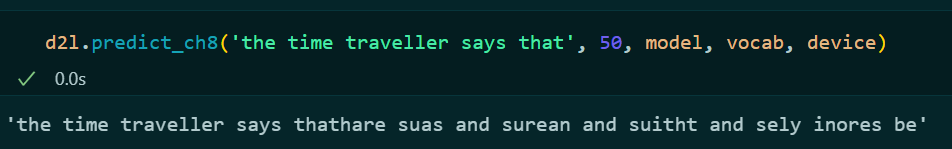
### 要求2

提示词为：traveller



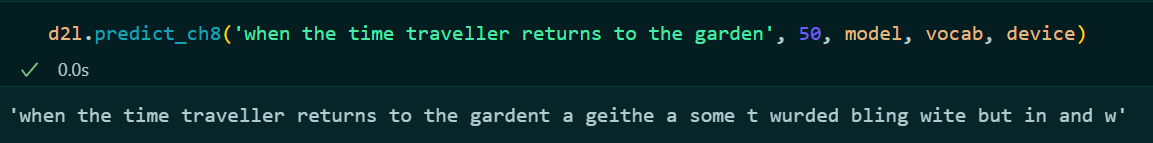
### 要求3

提示词为：the time traveller says that

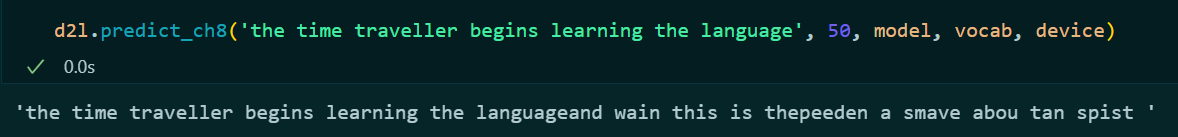


### 要求4

提示词为：when the time traveller returns to the garden

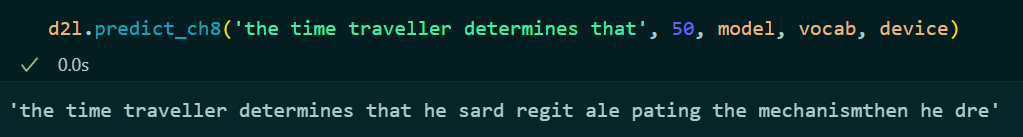


### 要求5

提示词为：the time traveller begins learning the language

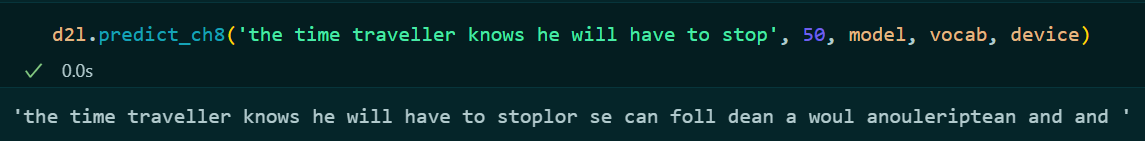
### 要求6

提示词为：the time traveller determines that



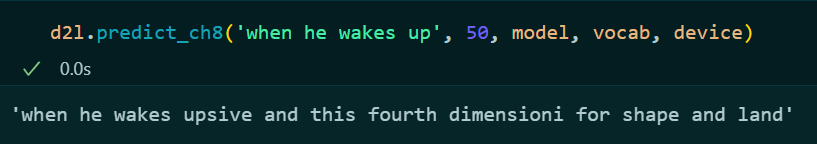
### 要求7

提示词为：the time traveller knows he will have to stop



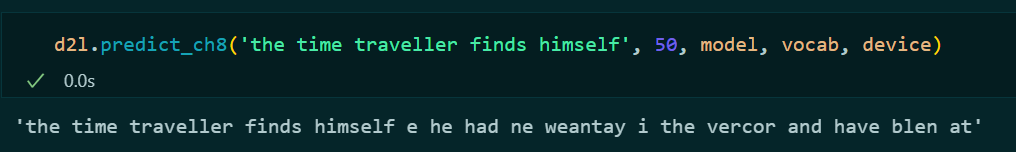
### 要求8

提示词为：when he wakes up



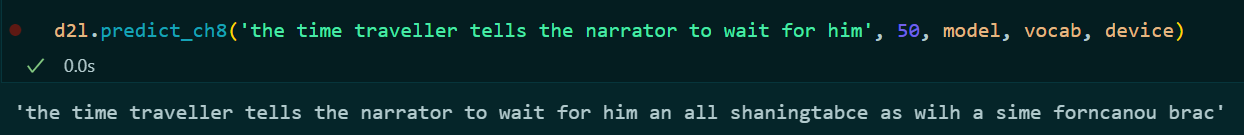
### 要求9

提示词为：the time traveller finds himself



### 要求10

提示词为：the time traveller tells the narrator to wait for him

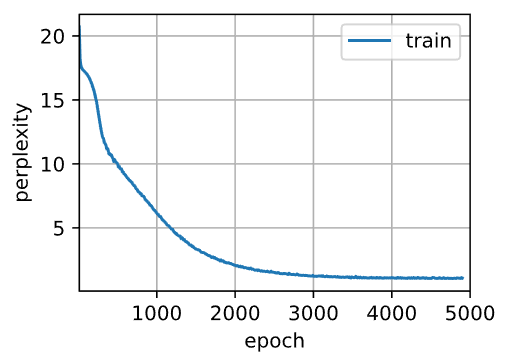


## 2. 附加题

### 要求1

修改参数与模型，使得输出更合理。

首先考虑减小学习率：将lr 从1变为0.1，同时将epoch从500增加到5000，运行结果为：

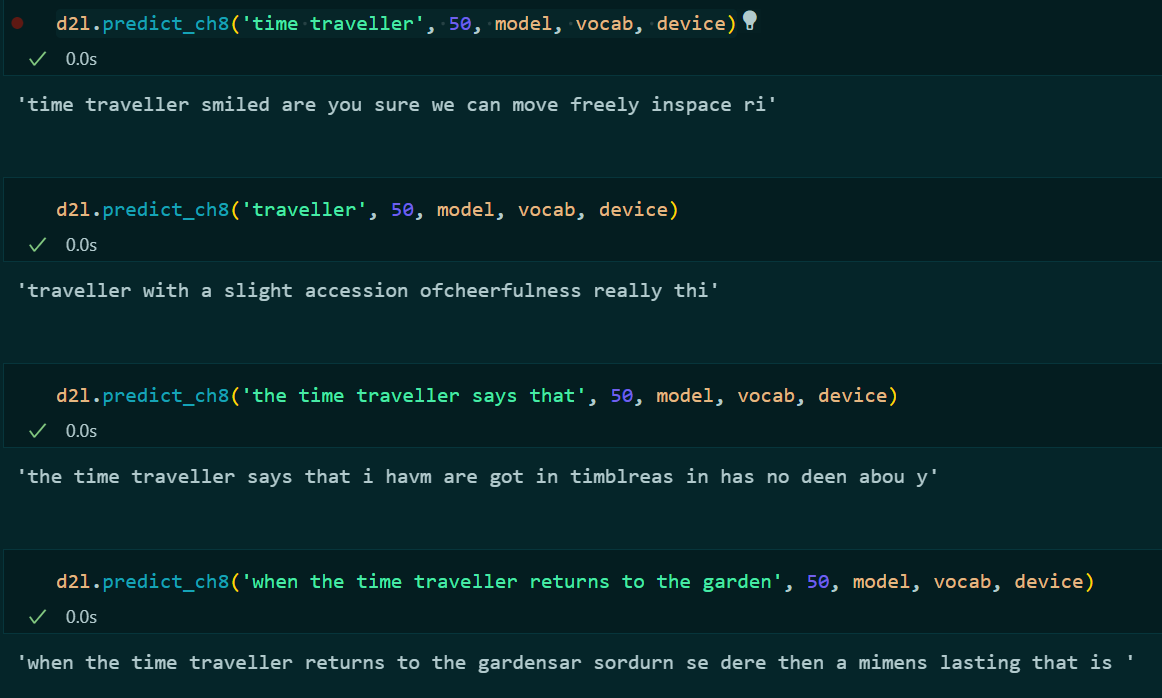


然而最终的困惑度为1.1，与原来参数的1.0相比，相差不大。

接下来考虑将batch\_size, num\_steps均增大为原来的2倍：

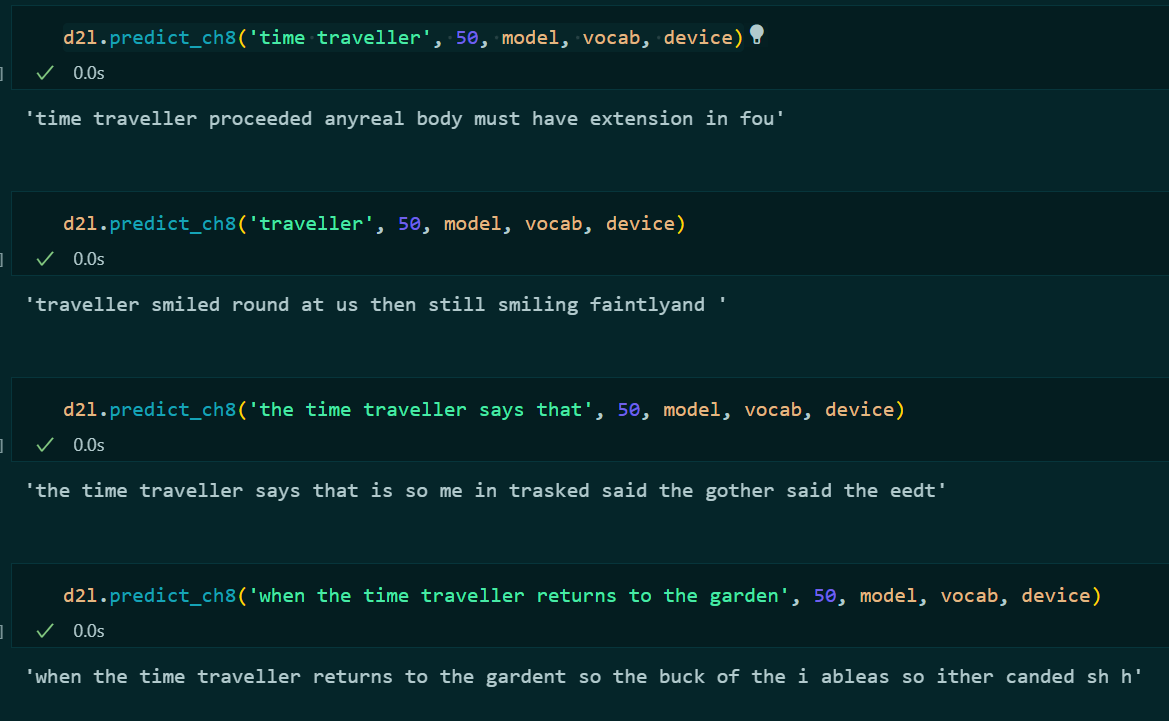
batch\_size**,** num\_steps **=** 64**,** 70

相应地，epoch增大到2000，运行结果：



可见，语义通顺度有了一定提升。

再考虑增大num\_hiddens的个数：从256更改为512：



可见，效果还不如之前的。所以将num\_hiddens仍然改为256.