

## 深度学习实验四-机器翻译

实验内容

数据集处理方法和实现方案

数据集处理函数实现

数据处理核心类 `DataLoader`

分批类 `Batch`

transformer模型和构建

整体结构

Encoder, Decoder

transformer实现

模型训练和测试说明

数据集处理

可执行文件说明

main.ipynb运行说明

main\_evaluate.ipynb运行说明

提交目录包含文件说明

未提交模型以及其他文件链接

测试集前50条数据测试结果对比

# 深度学习实验四-机器翻译

## 实验内容

复现transformer模型，实现机器翻译

## 数据集处理方法和实现方案

主要实现方法见文件 `lib` 文件夹下的 `DataLoader.py` 文件

## 数据集处理函数实现

### 数据处理核心类 `DataLoader`

- `DataLoader` 类实现了一下几个函数:

```
def __init__(self)
def load(self, en_path, cn_path, max_seq = 10000)
def build_dict(self, flag, kind, sentences, max_words = 1024 * 128)
def wordToID(self, en, cn, en_dict, cn_dict, sort = False)
def split_Batch(self, en, cn, batch_size, shuffle = True)
def prepare_train(self, g_max_seq1, g_max_seq2)
def prepare_test(self, g_max_seq3, test_path = args.testPath)
def prepare_train_data(self, train_path = args.trainPath, valid_path =
args.validPath)
def prepare_word_dict(self)
def prepare_train_id(self)
def prepare_train_batch(self)
def load_word_dict(self)
```

- 现对上述实现的函数做简要说明:

- `_init__`: 初始化函数
- `load(self, en_path, cn_path, max_seq = 10000)`: 数据集加载

参数说明:

**en\_path**: 英文数据集路径

**cn\_oath**: 中文数据集路径

**max\_seq**: 读取数据集的数据个数(由于设备的问题, 所有数据集无法全部加载)

返回值:

**en, cn**: 分词后的英文数据集和中文数据集

- `build_dict`: 词典构造

参数说明:

**flag**: **flag**为True时, 读取保存的字典, 否则重新生成字典

**kind**: 区分英文数据或是中文数据

**sentences**: 用于构建词典的句子

**max\_words**: 字典容纳最大的词的数量

返回值:

**word\_dict, total\_words, index\_dict**: 字典(词-索引), 字典中总词数, 字典(索引-词)

额外说明:

数据集处理时, 额外添加了四个词:

**PAD, BOS, EOS, UNK** = 'PAD', 'BOS', 'EOS', 'UNK'

用于表示填充字符, 句首标识符, 句尾标识符和不可识别字符。

此外, 当**flag**为False时, 将会把生成的字典保存下来, 方便后续调用。

保存的文件即:

- \* en\_index\_dic.npy
- \* en\_total\_words.txt
- \* en\_word\_dic.npy
- \* cn\_index\_dic.npy
- \* cn\_total\_words.txt
- \* cn\_word\_dic.npy

- `wordToID`: 句子序列化

参数说明:

**en**: 英文数据

**cn**: 中文数据

**en\_dict**: 英文字典

**cn\_dict**: 中文字典

**sort**: 是否排序, 默认不排序

- `split_Batch`: 数据集分批

参数说明:

**en**: 英文数据

**cn**: 中文数据

**batch\_size**: 批数据个数

**shuffle**: 是否打乱顺序

返回值:

分批后的数据集

- `prepare_train`: 准备训练使用的数据, 字典等各种参数

参数说明:

`g_max_seq1`: 训练集数据使用的个数

`g_max_seq2`: 验证集数据使用的个数

- `prepare_test`: 准备测试使用的数据, 字典等各种参数

参数说明:

`g_max_seq3`: 测试集数据使用的个数

`test_path`: 测试集数据路径

- `prepare_train_data`: 准备训练时所用的数据

参数说明:

`train_path`: 训练集数据路径

`valid_path`: 验证集数据路径

- `prepare_word_dict`: 调用生成字典函数, 生成训练时所用的字典
- `prepare_train_id`: 调用序列化函数, 将训练时的数据序列化
- `prepare_train_batch`: 调用数据分批函数, 将训练时的数据分批
- `load_word_dict`: 调用生成字典函数, 加载训练时产生的字典

## 分批类 Batch

通过对句子每个单词mask后,并分批,作为后续的网络输入.

# transformer模型和构建

## 整体结构

transformer模型主要分为两大部分:

- Encoder: 将自然语言序列映射为隐藏层
- Decoder: 再将隐藏层映射为自然语言序列 seq2seq + Attention的训练方式, 将注意力集中在解码端, transformer将注意力放在输入序列上, 对输入序列做Attention, 寻找序列内部的联系。

## Encoder, Decoder

关键则在实现一下几个功能类

- Embedding
- Positional Encoding:
- Attention
- Add & Normalize
- Feed forward
- Add & Normalize

在 `Mytransformer.py` 文件中对这几类进行了实现,详情请查看该文件.

## transformer实现

当前面所述的几个类完成以后,就可以根据论文描述的搭建起transformer类,具体请查看 `Mytransformer.py` 文件中的Transformer类.

```
class Transformer(nn.Module):
    def __init__(self, encoder, decoder, src_embed, tgt_embed, generator):
        super(Transformer, self).__init__()
        self.encoder = encoder
        self.decoder = decoder
        self.src_embed = src_embed
        self.tgt_embed = tgt_embed
        self.generator = generator

    def encode(self, src, src_mask):
        return self.encoder(self.src_embed(src), src_mask)

    def decode(self, memory, src_mask, tgt, tgt_mask):
        return self.decoder(self.tgt_embed(tgt), memory, src_mask, tgt_mask)

    def forward(self, src, tgt, src_mask, tgt_mask):
        return self.decode(self.encode(src, src_mask), src_mask, tgt, tgt_mask)
```

## 模型训练和测试说明

### 数据集处理

将数据集压缩后放在主目录的 `DataFolders` 文件夹下,即训练文件索引应 `./DataFolders/train_en`

### 可执行文件说明

文件提供了两个可以运行的notebook

- `main.ipynb`: 运行该文件即可训练文件,并调用训练好的模型进行测试
- `main_evaluate.ipynb`: 仅调用生成的模型进行预测

下面对该文件运行进行简要说明

### main.ipynb运行说明

实验运行平台为Google colab。

因此,不在该平台运行时,请注释文件的一下内容:

## 挂载

```
[ ] from google.colab import drive  
  
drive.mount('/content/drive')
```

## 切换位置

```
 %cd drive/MyDrive/Project4  
# !ls  
import sys  
sys.path.append('/content/drive/MyDrive/Project4')
```

需要注释的内容

文件中设置的参数如下:

```
set_seed(10)  
args.layers = 2  
args.batch_size = 128  
args.d_model = 256  
args.d_ff = 1024  
args.h_num = 8  
args.dropout = 0.1  
args.epochs = 20  
g_max_seq1 = 1500000 # 训练集数据使用前150w个  
g_max_seq2 = 150000 # 验证集数据使用前15w个  
g_max_seq3 = 100 # 测试前100个数据
```

这里补充说明一下, 由于测试过程缓慢, 因此对所有数据的测试单独设置为一个可执行文件, 这个将会在后面说明。

运行该文件即可训练模型, 训练好的模型保存为model.pkl文件。

过程会生成一些字典文件, 用于后续单独测试时使用。

## main\_evaluate.ipynb运行说明

运行该文件即可, 该文件将会对测试集所有文件进行测试。

所有预测的结果将会保存为result.txt文件。

过程中预测情况如下, 最终预测文件见文件链接。

```
测试样本: 19400,      bleu1: 0.6902  bleu2: 0.4001  bleu3: 0.3009  bleu4: 0.2941
19400 sentences has been translated!
测试样本: 19500,      bleu1: 0.6667  bleu2: 0.5      bleu3: 0.3293  bleu4: 0.1562
19500 sentences has been translated!
测试样本: 19600,      bleu1: 0.474   bleu2: 0.2505  bleu3: 0.1297  bleu4: 0.0532
19600 sentences has been translated!
测试样本: 19700,      bleu1: 0.2593  bleu2: 0.173   bleu3: 0.1062  bleu4: 0.0473
19700 sentences has been translated!
测试样本: 19800,      bleu1: 0.7667  bleu2: 0.6504  bleu3: 0.5813  bleu4: 0.5318
19800 sentences has been translated!
测试样本: 19900,      bleu1: 0.4283  bleu2: 0.265   bleu3: 0.1825  bleu4: 0.1292
19900 sentences has been translated!
测试样本: 20000,      bleu1: 0.4137  bleu2: 0.2697  bleu3: 0.1557  bleu4: 0.0564
20000 sentences has been translated!
测试样本: 20100,      bleu1: 0.4211  bleu2: 0.2163  bleu3: 0.065   bleu4: 0.0362
20100 sentences has been translated!
测试样本: 20200,      bleu1: 0.3542  bleu2: 0.2297  bleu3: 0.151   bleu4: 0.0935
20200 sentences has been translated!
测试样本: 20300,      bleu1: 0.3763  bleu2: 0.2916  bleu3: 0.2245  bleu4: 0.1669
20300 sentences has been translated!
测试样本: 20400,      bleu1: 0.5558  bleu2: 0.4557  bleu3: 0.3700  bleu4: 0.3340
20400 sentences has been translated!
```

预测过程示意

预测结束后，将会输出全部测试集平均的bleu值，如下所示:

```
测试样本: 401500,      bleu1: 0.255   bleu2: 0.1607  bleu3: 0.0975  bleu4: 0.0434
401500 sentences has been translated!
Bleu1: 0.417
Bleu2: 0.292
Bleu3: 0.201
Bleu4: 0.142
```

预测最终输出情况

因此训练的模型bleu值为:

Blue1	Bleu2	Bleu3	Bleu4
0.417	0.292	0.201	0.142

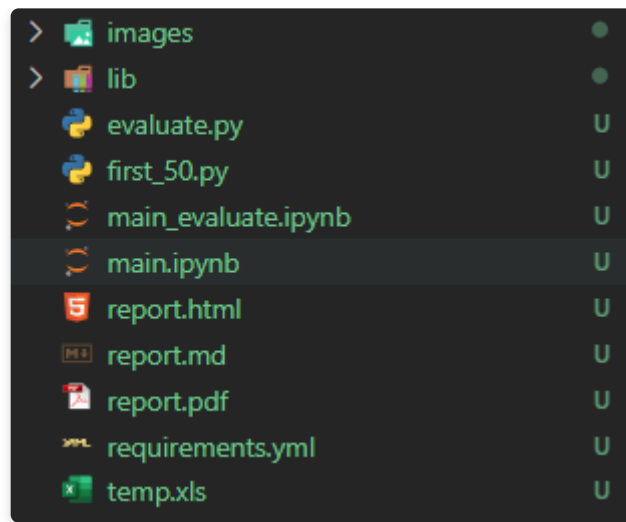
测试集前50条测试结果平均bleu情况如下，对比情况见文档末尾，全部测试集预测情况，请查看result.txt文件。

```
平均bleu值
Bleu1 : 0.4439 ; Bleu2 : 0.3164 ; Bleu3 : 0.2189 ; Bleu4 : 0.148
```

预测集前50条测试平均bleu值

## 提交目录包含文件说明

提交目录主要包含以下几个文件:



#### 提交目录

- images: 报告内容相关图片
- **lib文件夹**: 实现的模型，工具类等文件
- **evaluate.py**: 测试时调用的核心函数
- **first\_50.py**: 可执行文件，用于预测测试集的前50个句子，并将输出生成文件**temp.xls**
- **main\_evaluate.ipynb**: 可执行文件，用于调用模型预测**所有**测试集数据
- **main.ipynb**: 可执行文件，用于训练模型，并调用生成的模型进行预测
- **report.html, report.md, report.pdf**: 实验报告，可根据需要自由查看。
- **requirements.yml**: 依赖的第三方库
- **temp.xls**: 测试集前50句的对比情况。

## 未提交模型以及其他文件链接

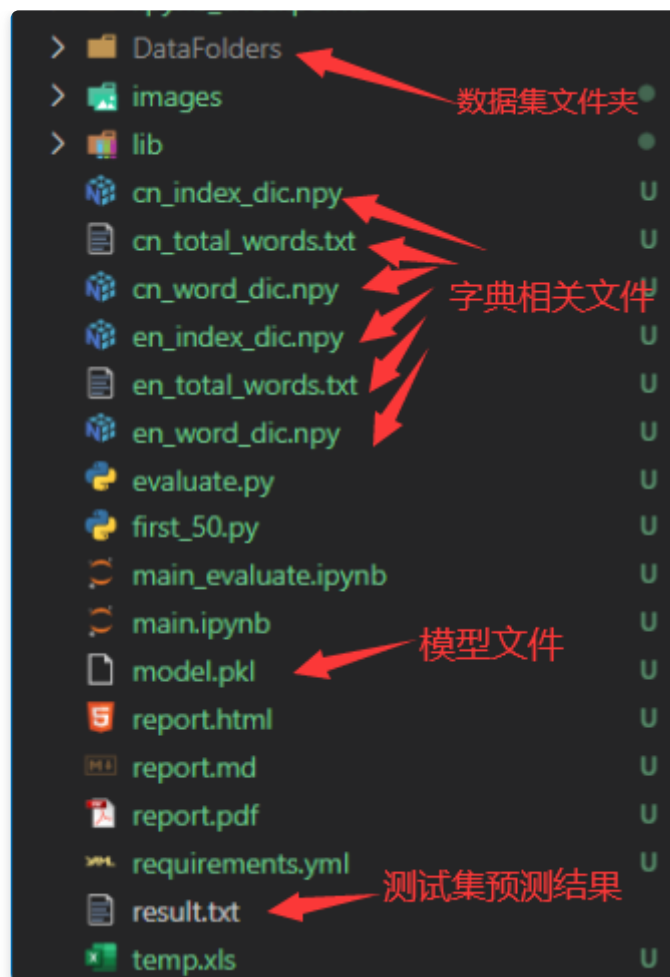
链接中主要包含以下几个文件

- 训练好的模型: **model.pkl**
- 数据集文件: **DataFolders**
- 通过训练数据集**生成的字典等文件**(也可以调用Data.prepare\_train(g\_max\_seq1, g\_max\_seq2))重新生成。

如果只参与测试，则需要使用链接中的文件

- 测试集预测结果文件: **result.txt**

使用方法：将文件压缩包解压后放在主目录下即可，文件结构应该为下图所示情况。



文件结构情况

文件链接: <https://pan.baidu.com/s/15MBXbBtx9MR8igT-q2-0Gg>

提取码: sdx

## 测试集前50条数据测试结果对比



id	label	value
0	英文原文	Results The key nursing methods were to promote nurses fundamental diathesis to carry out mental care for the patients and the nursing during acetazolamide stress brain perfusion SPECT.
	参考翻译	结果提高护理人员的基本素质；做好患者的心理护理和乙酰唑胺负荷脑血流灌注显像中的护理是关键性的护理措施。
	候选翻译	结果：提高护理人员的基本素质，为临床患者及护理工作者的脑灌注提供了依据。
	Bleu	Bleu1 : 0.4029 ; Bleu2 : 0.3374 ; Bleu3 : 0.2708 ; Bleu4 : 0.2374
1	英文原文	Notwithstanding his youth, Master Andrea was a very skilful and intelligent boy.
	参考翻译	另一方面, 纠正错误的浪费时间和抑制学生的积极性却是不言而喻的.
	候选翻译	尽管他年青年, 安德丽雅师长是一个很技巧的聪聪聪明的男孩。
	Bleu	Bleu1 : 0.1283 ; Bleu2 : 0.0207 ; Bleu3 : 0.0114 ; Bleu4 : 0.0085
2	英文原文	Worryingly, with only weeks before the law comes into effect, it has not yet been announced who will oversee it.
	参考翻译	更令人焦虑的是，距这部法律实施已经没有几周时间了，但是现在还没有宣布谁来监督这部法律。
	候选翻译	令人担心的是，只有几周之前才会出现，还没有宣布监督谁监督。
	Bleu	Bleu1 : 0.4043 ; Bleu2 : 0.2985 ; Bleu3 : 0.2167 ; Bleu4 : 0.1483
3	英文原文	The impact of the car hitting the tree killed the driver.
	参考翻译	司机死于汽车撞到树后产生的冲击力.
	候选翻译	车撞撞撞撞车撞死了司机的撞击撞击死了司机。
	Bleu	Bleu1 : 0.3182 ; Bleu2 : 0.1741 ; Bleu3 : 0.0533 ; Bleu4 : 0.0299
4	英文原文	The study of contacts of rough surfaces is a young and interdisciplinary field.
	参考翻译	粗糙表面接触研究是一门年轻的交叉学科。
	候选翻译	粗糙表面接触的研究是一个年幼的跨学科领域，是一个跨学科领域。
	Bleu	Bleu1 : 0.5 ; Bleu2 : 0.3939 ; Bleu3 : 0.3216 ; Bleu4 : 0.2649

5	英文原文	If you know someone who is a consultant, you should talk to that person.
	参考翻译	如果你知道有人是顾问，你应该找人。
	候选翻译	如果你知道一个顾问的人，你应该和那个人交谈。
	Bleu	Bleu1 : 0.6364 ; Bleu2 : 0.4924 ; Bleu3 : 0.3928 ; Bleu4 : 0.3128
6	英文原文	Some plants flower year in and year out.
	参考翻译	有些植物一年到头都开花。
	候选翻译	一些植花年年年年花花花花。
	Bleu	Bleu1 : 0.3571 ; Bleu2 : 0.1657 ; Bleu3 : 0.0612 ; Bleu4 : 0.038
7	英文原文	Demographic and clinical characteristics of individuals in a bipolar disorder case registry.
	参考翻译	双相障碍档案上的人口统计的特点和临床特征。
	候选翻译	双双极性疾病患病症状病症状表中个人的人人人口统和临床特征。
	Bleu	Bleu1 : 0.3667 ; Bleu2 : 0.318 ; Bleu3 : 0.2624 ; Bleu4 : 0.2117
8	英文原文	One of the program's priorities is to combine family and work life and develop more flexible and effective forms of work organization and support services.
	参考翻译	该项目的一个优先事项就是兼顾家庭与工作生活，制订更加灵活和更加有效的工作组织形式和支助服务形式。
	候选翻译	其中一个项目的优先事项是把家庭和工作生活结合起来，发展更灵活、更有效的工作组织和支持服务。
	Bleu	Bleu1 : 0.7068 ; Bleu2 : 0.5344 ; Bleu3 : 0.396 ; Bleu4 : 0.3019
9	英文原文	Both silhouette preserving and shading preserving criteria are satisfied by applying a new error metric-constrained normal cone to view-independent simplification.
	参考翻译	算法为地表的简化提出一种新的基于受限法向锥的误差计算方法，使得模型简化具有轮廓保持和光照保持特性。
	候选翻译	将一个新的误差锥形图保护和遮保标准均满足保护条件的保护，以便简化。
	Bleu	Bleu1 : 0.2426 ; Bleu2 : 0.1366 ; Bleu3 : 0.0718 ; Bleu4 : 0.0295

10	英文原文	Information technology - Open Systems Interconnection - Connection-oriented presentation protocol: Protocol specification
	参考翻译	信息技术.开放系统互连.面向连接型表示协议: 协议规范
	候选翻译	信息技开开系统互联网.提交协议议议协议.协议规范.协议规规范
	Bleu	Bleu1 : 0.5 ; Bleu2 : 0.3714 ; Bleu3 : 0.2701 ; Bleu4 : 0.1644
11	英文原文	Alan Beattie is the World Trade Editor of the Financial Times, leading the paper's coverage of trade policy and economic globalisation.
	参考翻译	艾伦·贝蒂是金融时报的世贸编辑, 负责领导该报覆盖有关贸易政策与经济全球化的报道。
	候选翻译	艾伦·诗是金融时代的世贸编辑, 导论文对贸易政策和经济全球化的报道。
	Bleu	Bleu1 : 0.6373 ; Bleu2 : 0.5676 ; Bleu3 : 0.5015 ; Bleu4 : 0.4398
12	英文原文	She kind of hoped he would come.
	参考翻译	她多少有点希望他会来。
	候选翻译	她希望他能来。她希望他能来。
	Bleu	Bleu1 : 0.4286 ; Bleu2 : 0.3145 ; Bleu3 : 0.202 ; Bleu4 : 0.093
13	英文原文	In oil seeds they may occupy a large fraction of total volume of reserve cells.
	参考翻译	在油质种子中,它们可占据贮藏细胞总体积的大部分。
	候选翻译	在油种中, 他们可占占有大部分储备细胞总体积的总数。
	Bleu	Bleu1 : 0.64 ; Bleu2 : 0.5164 ; Bleu3 : 0.4113 ; Bleu4 : 0.3121
14	英文原文	The source of the starting material need not be identified, but may be requested.
	参考翻译	起始原料的来源通常无需说明, 但有时会要。
	候选翻译	开始材料的源源不需要确认, 但可能需要请求。
	Bleu	Bleu1 : 0.3333 ; Bleu2 : 0.1291 ; Bleu3 : 0.0444 ; Bleu4 : 0.0264
15	英文原文	He'll hit the ceiling if I ask to change again.

	参考翻译	如果我要求再改变的话他会气疯的。
	候选翻译	如果我再次改变，他就打天天花板，我就要打上天花板。
	Bleu	Bleu1 : 0.36 ; Bleu2 : 0.2121 ; Bleu3 : 0.1251 ; Bleu4 : 0.0546
16	英文原文	Mao Zedong's agricultural economic thought is through the whole process of Chinese revolution and construction .
	参考翻译	毛泽东的农业经济思想贯穿着中国革命和建设的全过程。
	候选翻译	毛泽东农业经济思想是中国革命和建设的全过程。
	Bleu	Bleu1 : 0.68 ; Bleu2 : 0.4761 ; Bleu3 : 0.3092 ; Bleu4 : 0.1077
17	英文原文	Comprehensive multi-level exploration of buried active fault: an example of Yinchuan buried active fault
	参考翻译	隐伏活动断层的多层次综合探测&以银川隐伏活动断层为例
	候选翻译	综合综合地理活故障综合综合多层探&以银川隐积活故障为例
	Bleu	Bleu1 : 0.5 ; Bleu2 : 0.36 ; Bleu3 : 0.2464 ; Bleu4 : 0.186
18	英文原文	Third the process of fighting for the leadership we must pay attention to the party and the people.
	参考翻译	第三，在争取领导权过程中对党和人民群众的重视。
	候选翻译	第三，我们必须注重党人和人民的领导进程。
	Bleu	Bleu1 : 0.5322 ; Bleu2 : 0.3386 ; Bleu3 : 0.2185 ; Bleu4 : 0.0842
19	英文原文	Analysis of the protection of all-bridged GTR resistance welding inverter
	参考翻译	全桥式GTR逆变电阻焊电源的保护分析
	候选翻译	引引引引焊耐焊逆变逆焊焊逆变器防护的分析分析分析了引焊耐焊逆变逆变器的防护性分析分析。
	Bleu	Bleu1 : 0.1628 ; Bleu2 : 0.088 ; Bleu3 : 0.0266 ; Bleu4 : 0.0147
20	英文原文	The purpose of my thesis is to find out the suitable credit risk measurement model in our country, which can improve the credit risk management level of our banks.
	参考翻译	本文研究的主旨在于通过分析和比较现代信用风险度量模型，试图找出适合我国实际的信用风险模型，以提高我国银行业的竞争力。

	候选翻译	本文旨在探索我国的适当信用风险评估模型，可以提高我国银行信用卡风险管理水平。
	Bleu	Bleu1 : 0.4137 ; Bleu2 : 0.3328 ; Bleu3 : 0.2618 ; Bleu4 : 0.2058
21	英文原文	Ben's hobby is collecting stamps. He has many beautiful stamps. He is showing them to his classmates.
	参考翻译	本的业余爱好是集邮。他有许多漂亮的邮票。他正在把它们给他的同学们看。
	候选翻译	本的爱好是集邮邮邮。他有很多美丽的邮票，他正在向同学展示他们。
	Bleu	Bleu1 : 0.6442 ; Bleu2 : 0.5034 ; Bleu3 : 0.3815 ; Bleu4 : 0.2711
22	英文原文	While travelling, they entrusted their children to the care of baby sitter.
	参考翻译	他们在若干名士兵的保护下外出旅行。
	候选翻译	旅行时，他们把孩子们的孩子送给宝宝保护照护保护保护。
	Bleu	Bleu1 : 0.2692 ; Bleu2 : 0.1797 ; Bleu3 : 0.0513 ; Bleu4 : 0.0277
23	英文原文	Pour into the prepared baking pan.
	参考翻译	倒入备好的玻璃烤盘内。
	候选翻译	倒入准备的烤盘里，倒入做烤盘。
	Bleu	Bleu1 : 0.4667 ; Bleu2 : 0.2582 ; Bleu3 : 0.08 ; Bleu4 : 0.0455
24	英文原文	Meanwhile, China's political cycle may exacerbate risks of an asset bubble.
	参考翻译	同时，中国的政治周期也可能加剧资产泡沫风险。
	候选翻译	同时，中国政循环可能会加剧资产泡泡泡泡泡沫的风险。
	Bleu	Bleu1 : 0.72 ; Bleu2 : 0.6 ; Bleu3 : 0.4785 ; Bleu4 : 0.3757
25	英文原文	It was in light of the above spirit that China took a constructive part in the Security Council discussions. We are ready to stay in communication with relevant parties over this.
	参考翻译	中方秉承上述精神，以建设性的态度参加了安理会有关讨论，我们愿就此与有关各方继续保持沟通。
	候选翻译	鉴于上述精神，中国在安全理事会讨论中取得建设性的成果，我们准备与有关的各方进行沟通。

	Bleu	Bleu1 : 0.5903 ; Bleu2 : 0.4538 ; Bleu3 : 0.3399 ; Bleu4 : 0.2317
26	英文原文	Many put plenty of stock in a team's success.
	参考翻译	很多人将球队的战绩看得很重。
	候选翻译	许多人把大量的股票投入到团队成功中。
	Bleu	Bleu1 : 0.2381 ; Bleu2 : 0.1091 ; Bleu3 : 0.0397 ; Bleu4 : 0.0243
27	英文原文	In his mind a voice cried, Now is the time of testing.
	参考翻译	他脑海中有个声音在大喊, 现在正是一决雌雄的时候了.
	候选翻译	在他的脑声里哭了, 现在是测试的时候。
	Bleu	Bleu1 : 0.4142 ; Bleu2 : 0.2226 ; Bleu3 : 0.128 ; Bleu4 : 0.0555
28	英文原文	These patients also were receiving drugs known as azathioprine or 6 - mercaptopurine.
	参考翻译	这些患者同时接受了硫唑嘌呤或巯基嘌呤治疗.
	候选翻译	这些病人也被称为硫唑嘌呤或6或6例药物。
	Bleu	Bleu1 : 0.2378 ; Bleu2 : 0.1543 ; Bleu3 : 0.0501 ; Bleu4 : 0.029
29	英文原文	The use of calcium sulfate requires an greater capital investment .
	参考翻译	使用硫酸钙做原料, 需要更多的基建投资。
	候选翻译	硫硫酸钙的使用需要更大的资本投资, 需要更大的资本投资。
	Bleu	Bleu1 : 0.4815 ; Bleu2 : 0.3849 ; Bleu3 : 0.2873 ; Bleu4 : 0.1773
30	英文原文	Mice doing obeisance for peach
	参考翻译	景点名称: 老鼠拜桃
	候选翻译	老鼠为桃子做客守守守守桃守桃守的守桃子。桃子的守守守守守。桃子桃。
	Bleu	Bleu1 : 0.0909 ; Bleu2 : 0.0533 ; Bleu3 : 0.0209 ; Bleu4 : 0.0132

31	英文原文	In the paper analysis of wear mechanism of circular diamond saw blades with several typical segment structures being widely used in stone processing area.
	参考翻译	本文对目前石材加工领域广泛使用的具有几种典型节块结构的金刚石圆锯片的磨损机理进行了一些分析。
	候选翻译	本文分析了圆钻石锯刀的磨损机理，具有几种典型的段结构在石加工领域广泛应用。
	Bleu	Bleu1 : 0.6569 ; Bleu2 : 0.4932 ; Bleu3 : 0.3914 ; Bleu4 : 0.3243
32	英文原文	How did leaving Portugal for Spain affect you development as a coach?
	参考翻译	离开葡萄牙去西班牙对你发展成一名教练有什么影响？
	候选翻译	西西西西班牙如何影响你的发展作为教教练的影响你们的发展？如何？
	Bleu	Bleu1 : 0.3125 ; Bleu2 : 0.2245 ; Bleu3 : 0.1189 ; Bleu4 : 0.0491
33	英文原文	Object used to create a new child node at the beginning of the list of child nodes of the current node.
	参考翻译	对象,该对象用于在当前节点的子节点列表的开始处创建一个新的子节点。
	候选翻译	用于在当当节点的初创新节点的子节点列表中创建新的儿节节点。
	Bleu	Bleu1 : 0.6032 ; Bleu2 : 0.5133 ; Bleu3 : 0.4248 ; Bleu4 : 0.3523
34	英文原文	The calculation results of unsteady flow in an exhaust manifold model are presented.
	参考翻译	最后对一根模型排气管进行了不稳定的流场计算。
	候选翻译	提出了排气管流模型中不稳定流流流的计算结果。
	Bleu	Bleu1 : 0.6364 ; Bleu2 : 0.4264 ; Bleu3 : 0.263 ; Bleu4 : 0.0989
35	英文原文	To avoid these negative effects bring a false academic prosperity, academic theorists have already explored the issue too much.
	参考翻译	为尽可能地避免这些负面影响给高校学术带来虚假繁荣，理论界已经对高校学术异化问题进行了许多的探索。
	候选翻译	为避免这些负面影响导致虚假学术繁荣，学理论家已经探究了这一问题。
	Bleu	Bleu1 : 0.4928 ; Bleu2 : 0.3674 ; Bleu3 : 0.2673 ; Bleu4 : 0.2114
36	英文原文	The design route deviation succinct isgraceful, the vitality long-time clothing to the manufacture.

	参考 翻译	设计路线偏向于制作简洁优雅、生命力长久的服饰。
	候选 翻译	设计路线偏简洁，长时间服装对制造工业的生命力长。
	Bleu	Bleu1 : 0.5385 ; Bleu2 : 0.3595 ; Bleu3 : 0.2528 ; Bleu4 : 0.1628
37	英文 原文	The company doubled up with hilarity.
	参考 翻译	大伙儿乐得弯下腰来。
	候选 翻译	公司狂欢地使公司变翻了一倍，狂欢地大笑。
	Bleu	Bleu1 : 0.0476 ; Bleu2 : 0.0154 ; Bleu3 : 0.0108 ; Bleu4 : 0.0091
38	英文 原文	The rock left a chip in the paint.
	参考 翻译	石头在油漆留下了一道缺口。
	候选 翻译	石石石在油漆上留了一块芯片。
	Bleu	Bleu1 : 0.5 ; Bleu2 : 0.3397 ; Bleu3 : 0.2126 ; Bleu4 : 0.0967
39	英文 原文	Wish you all A Fantastic Happy Rabbit Year 2011.
	参考 翻译	祝大家兔年快乐，福星高照。恭喜恭喜！
	候选 翻译	祝你们全一个美好快快兔年一年一年——美好的快乐兔子。
	Bleu	Bleu1 : 0.2222 ; Bleu2 : 0.1307 ; Bleu3 : 0.0409 ; Bleu4 : 0.0231
40	英文 原文	Do you, Princeton Girl, feel like you made the right choice meeting me here tonight?
	参考 翻译	普林斯顿女孩儿，你觉得今天晚上来见我的决定做的对吗？
	候选 翻译	你是吗，普林斯林顿女孩，你觉得你今晚在这里做出了正确的选择会我的选择吗？
	Bleu	Bleu1 : 0.4722 ; Bleu2 : 0.3285 ; Bleu3 : 0.2333 ; Bleu4 : 0.1401
41	英文 原文	eat, drink, and be merry -- that's his philosophy
	参考 翻译	吃喝玩乐,那就是他的人生哲学



	候选翻译	吃吃喝酒，喝快乐--那是他的哲哲哲学。
	Bleu	Bleu1 : 0.4737 ; Bleu2 : 0.3244 ; Bleu3 : 0.1836 ; Bleu4 : 0.0789
42	英文原文	He was able to make other noises , for instance , short , cough-like grunts , accompanying his pantomimed communications .
	参考翻译	他可以发出噪音，例如短暂、像是咳嗽呼噜声并且伴随著手势上的沟通。
	候选翻译	例如，他能够制造噪音，短短，呼声呼声，伴随他的通信。
	Bleu	Bleu1 : 0.4275 ; Bleu2 : 0.233 ; Bleu3 : 0.1216 ; Bleu4 : 0.0499
43	英文原文	on the other hand, we found a new and easier way to reconstruct nuclei in the central nervous system.
	参考翻译	另一方面找到了一种较简单易行的立体重建中枢神经系统神经核团的新方法。
	候选翻译	另一方面，我们发现了一种新的方式，在中枢神经系统中重构核核的新方法。
	Bleu	Bleu1 : 0.6176 ; Bleu2 : 0.5119 ; Bleu3 : 0.4342 ; Bleu4 : 0.3548
44	英文原文	The absence of it is not an assured ground of condemnation, but the presence of it is an invariable sign of goodness of heart.
	参考翻译	它的缺失虽然不能成为你被谴责的理由，但是存有爱美之心却是拥有一颗善良之心的不变的标志。
	候选翻译	没有被放心的谴责，但是它的存在是一个不变的迹象。
	Bleu	Bleu1 : 0.3398 ; Bleu2 : 0.2165 ; Bleu3 : 0.1245 ; Bleu4 : 0.0452
45	英文原文	We need not have hurried.
	参考翻译	我们当时实在不必那么匆忙。
	候选翻译	我们不需急急急急急急。
	Bleu	Bleu1 : 0.23 ; Bleu2 : 0.1387 ; Bleu3 : 0.0561 ; Bleu4 : 0.0367
46	英文原文	Return result sets that include many rows, from several thousand to several million rows.
	参考翻译	返回包括很多行的结果集，从数千行到数百万行。
	候选翻译	返结结结果集包括许多行，从数千万到几千万行。
	Bleu	Bleu1 : 0.7273 ; Bleu2 : 0.5583 ; Bleu3 : 0.3965 ; Bleu4 : 0.2393

47	英文原文	EXAMPLE " Whoever wins in Florida will have enough votes to break the tie. "
	参考翻译	谁赢了弗 罗里 达谁就能打破僵局获得足够的选票赢得比赛。
	候选翻译	在佛罗佛达州赢的人有足够的选票可以打破这场比赛。
	Bleu	Bleu1 : 0.46 ; Bleu2 : 0.3323 ; Bleu3 : 0.2402 ; Bleu4 : 0.1867
48	英文原文	The judge had made a terrible bluder, so he had to step down.
	参考翻译	法官犯了大错，所以他只好下台。
	候选翻译	法官犯了一个可怕的事，所以他必须下台。
	Bleu	Bleu1 : 0.5789 ; Bleu2 : 0.5073 ; Bleu3 : 0.423 ; Bleu4 : 0.3119
49	英文原文	That colored one is over beautiful.
	参考翻译	那个彩色的很漂亮。
	候选翻译	那彩色的人是美丽的美。
	Bleu	Bleu1 : 0.4545 ; Bleu2 : 0.3015 ; Bleu3 : 0.2162 ; Bleu4 : 0.106