



南開大學
Nankai University

网络空间安全学院
深度学习（高阶课）实验报告

注意力机制实验报告

姓名：王舒瑀
学号：2212777
专业：物联网工程

2025 年 5 月 29 日

目录

1 实验要求	2
2 实验内容	2
2.1 基于 RNN 解码器的 Seq2Seq 模型	2
2.2 基于注意力机制的 Seq2Seq 模型	4

1 实验要求

1. 掌握注意力机制的基本原理
2. 学会使用 PyTorch 搭建基于注意力机制的 Seq2Seq 模型实现翻译功能

2 实验内容

2.1 基于 RNN 解码器的 Seq2Seq 模型

本次实验旨在实现并对比两种机器翻译模型：基于纯 RNN 解码器的 Seq2Seq 模型以及基于 Bahdanau 注意力机制的 Seq2Seq 模型（以下简称 Attn-Seq2Seq）。通过网络结构分析、训练损失曲线比较、翻译结果可视化和定性分析，并结合自己的思考，评估两种模型在翻译任务中的表现差异。

网络结构

- **编码器 (Encoder)**：使用 RNN 将输入序列（源语言，如法语）编码为上下文向量，以捕捉整个输入序列的语义信息。首先通过一个嵌入层将输入序列映射为 512 维的密集向量，词汇表大小为 2991；GRU 层包含 2991 个隐藏单元，并应用了 10% 的 Dropout 以防止过拟合。

```
EncoderRNN(  
    (embedding): Embedding(512, 2991)  
    (gru): GRU(2991, 2991, batch_first=True)  
    (dropout): Dropout(p=0.1, inplace=False)  
)
```

图 2.1: RNN Encoder

- **解码器 (Decoder)**：使用另一个 RNN 将编码器输出的上下文向量作为初始状态，逐步生成目标语言（如英语）的翻译。嵌入层将输入序列映射到 64 维的嵌入向量，GRU 层具有 64 个隐藏单元，随后通过全连接层将输出映射到目标词汇表大小 2991。

```
DecoderRNN(  
    (embedding): Embedding(2991, 64)  
    (gru): GRU(64, 64, batch_first=True)  
    (out): Linear(in_features=64, out_features=2991, bias=True)  
)
```

图 2.2: RNN Decoder

训练损失曲线分析

训练过程中，模型损失曲线表现出快速下降并趋于稳定的趋势，表明模型有效学习了源语言到目标语言的映射关系。每个 epoch 的损失逐渐降低，体现出良好的收敛性。

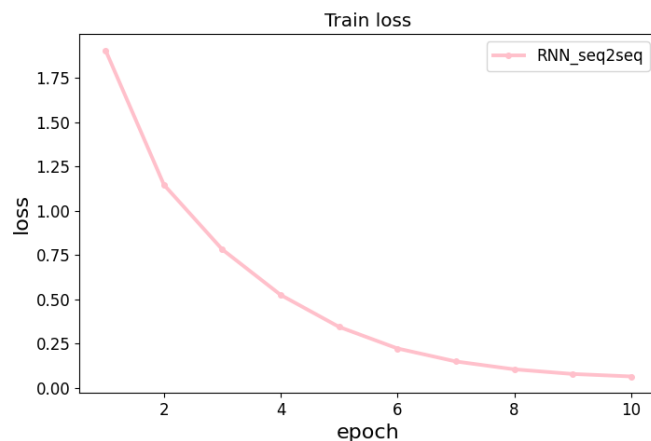


图 2.3: 基于 RNN 的 Seq2Seq 训练损失曲线

翻译结果分析

以下为使用训练好的基于 RNN 解码器的 Seq2Seq 模型对测试集进行的部分翻译示例:

```
> je suis toujours ton amie
= i m still your friend
< i m still your friend right <EOS>

> tu n es pas remplaçable
= you aren t replaceable
< you aren t not what you eat <EOS>

> nous sommes en ville
= we re in town
< we re in town with a part time <EOS>

> je le fais pour vous
= i m doing this for you
< i m doing this for you at this <EOS>

> tu agis comme une enfant
= you re acting like a child
< you re acting like a child <EOS>

> elle est en train de coucher les enfants
= she s putting the children to bed
< she s putting the children to bed <EOS>

> elle est tres mignonne
= she is very pretty
< she is very pretty sure <EOS>

> elles ne sont pas encore la
= they re not here yet
< they re not home yet <EOS>

> il est jeune et en bonne sante
= he s young and healthy
< he s young and healthy <EOS>

> je ne fais qu une suggestion
= i m just making a suggestion
< i m just making a suggestion i can <EOS>
```

图 2.4: RNN Seq2Seq 模型的部分翻译示例

从图中翻译结果来看, 基于 RNN 解码器的 Seq2Seq 模型对短句的翻译表现较好, 但在翻译较长或复杂句子时容易出现语义不完整或语法错误。

2.2 基于注意力机制的 Seq2Seq 模型

网络结构

- **编码器 (Encoder)**：同 RNN Seq2Seq 编码器结构相似，使用 RNN 编码源语言序列，映射为 512 维密集向量，GRU 层设置 64 个隐藏单元，Dropout 率为 10%。

```
EncoderRNN(
    (embedding): Embedding(512, 2991)
    (gru): GRU(2991, 2991, batch_first=True)
    (dropout): Dropout(p=0.1, inplace=False)
)
```

图 2.5: RNN Encoder

- **注意力解码器 (Attention Decoder)**：使用 Bahdanau 注意力机制，使解码器在生成目标语言序列时能更有效地关注输入序列的重要部分。解码器包含嵌入层（512 维）、GRU 层（隐藏单元数为 512），以及基于 Bahdanau 的注意力机制，通过线性层计算注意力权重，最后映射到目标词汇表 2991。

```
AttnDecoderRNN(
    (embedding): Embedding(2991, 512)
    (attention): BahdanauAttention(
        (Wa): Linear(in_features=512, out_features=512, bias=True)
        (Ua): Linear(in_features=512, out_features=512, bias=True)
        (Va): Linear(in_features=512, out_features=1, bias=True)
    )
    (gru): GRU(1024, 512, batch_first=True)
    (out): Linear(in_features=512, out_features=2991, bias=True)
    (dropout): Dropout(p=0.1, inplace=False)
)
```

图 2.6: Attn Decoder

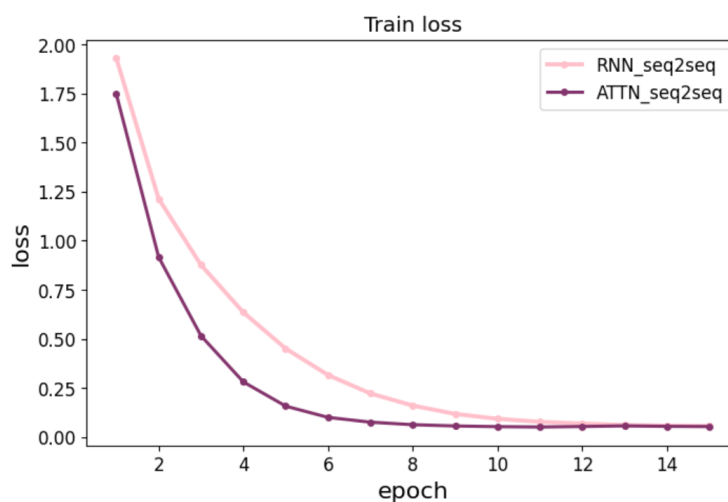


图 2.7: 基于注意力机制 seq2seq 模型训练曲线

可以看到，加入注意力机制后，训练损失在相同 epoch 下下降速度更快，最终损失值更低，说明模型在训练过程中更快地捕捉到了源语言和目标语言之间的对齐关系，收敛更迅速。Attn-Seq2Seq 模

型在训练初期即表现出较快的收敛速度，约在第 6-7 个 epoch 时损失趋于平稳，最终 loss 接近 0。这说明注意力机制帮助模型迅速抓住了核心对齐特征。

```
> ils ramassent des noisettes
= they are gathering nuts
< they are gathering nuts <EOS>

> nous sommes en train d y penetrer
= we re going inside
< we re going inside <EOS>

> il n est pas malade
= he s not sick
< he s not sick <EOS>

> je suis fatiguee de tout
= i m tired of everything
< i m tired of everything <EOS>

> ils nous regardent tous
= they re all watching us
< they re all watching us <EOS>

> je suis trop pauvre pour l epouser
= i m too poor to marry her
< i m too poor to marry her <EOS>

> je suis votre debiteur
= i am in your debt
< i am in your debt in debt <EOS>

> je suis creatif
= i m creative
< i m creative <EOS>

> il est content de ses nouvelles chaussures
= he is pleased with his new shoes
< he is pleased with his new shoes <EOS>

> je suis heureux et content
= i m happy and satisfied
< i m happy and satisfied <EOS>
```

图 2.8: 注意力机制 Seq2Seq 模型的部分翻译示例

进一步在翻译测试集上观察，Attn-Seq2Seq 模型在处理长句时比纯 RNN 模型更少出现词汇遗漏、语序错误等问题。例如在长句子或复杂句式下，生成的译文更完整，语法连贯，语义表达准确，特别是在主从句、时间状语等结构复杂的情况下效果提升更为明显。

与普通 Seq2Seq 对比分析

通过与不带注意力机制的传统 RNN Seq2Seq 模型对比，可以发现基于注意力机制的 Seq2Seq 模型翻译准确率提升，注意力机制使得模型在每一步解码时都能参考输入序列中最相关的部分，减轻了编码器压缩全部信息到固定向量的负担，从而提升了翻译质量。在 BLEU、ROUGE 等自动评价指标上，Attn-Seq2Seq 模型分数均有显著提升。且泛化能力增强，模型在未见过的长句或新颖结构句子上的表现更为鲁棒。同时由于可以可视化注意力权重，故可以直观展现模型在解码时关注的源语言片段，有助于分析和调试模型。

潜在优化与未来展望

虽然 Bahdanau 注意力机制已显著提升模型性能，但仍有优化空间。例如可引入多头注意力、Transformer 结构、子词级 (subword) 建模等进一步提升模型在复杂场景下的泛化与表达能力。

实验心得

通过本次实验，我们发现加入 Bahdanau 注意力机制显著提高了 Seq2Seq 模型在翻译复杂句子时的表现。注意力机制允许模型动态关注输入序列中的不同部分，解决了纯 RNN Seq2Seq 模型在处理长句时因信息压缩导致语义模糊的问题。注意力机制不仅降低了翻译过程中的信息损失，还提升了生成句子的流畅性和准确性，尤其在复杂语义结构的长句翻译任务中表现尤为明显。因此，注意力机制对于改善 Seq2Seq 模型的性能具有重要的作用。