

---

# Pie-Lab 2025年暑期培训

---

你的名字

Pie-Lab

北京理工大学计算机学院  
北京市海淀区中关村南大街  
786362411@qq.com

## Abstract

以下是IMDB数据集上的文本情感二分类任务的实验报告，包括网络结构的原理解释、模型结构设计与损失函数选择、实验设置、实验结果与可视化和总结与分析。

## 1 网络结构的原理解释

### 1.1 LSTM模型(Long Short-Term Memory)

在深度学习领域，循环神经网络（RNN）作为处理序列数据的利器，一直备受关注。然而，传统 RNN 在处理长序列数据时，容易出现梯度消失或梯度爆炸的问题，导致难以学习到长距离的依赖关系。为了解决这一难题，长短期记忆网络（Long Short-Term Memory, LSTM）应运而生。

LSTM 是一种特殊的 RNN，它通过引入“门控机制”，有效解决了传统 RNN 在处理长序列数据时的缺陷，能够更好地捕捉长距离依赖关系。在自然语言处理、时间序列预测、语音识别等众多领域，LSTM 都有着广泛且出色的表现。例如在自然语言处理中，它可以记住前文提到的关键信息，从而更好地理解和生成后续文本。

LSTM 网络的核心结构由细胞状态（Cell State）和三个门控单元组成，这三个门控单元分别是遗忘门（Forget Gate）、输入门（Input Gate）和输出门（Output Gate）。

#### 1. 细胞状态：

细胞状态贯穿整个 LSTM 网络，能够在序列的不同时间步之间传递信息，实现长期记忆的功能。信息在细胞状态上的传递相对稳定，不会像传统 RNN 那样在传递过程中因梯度问题而丢失信息。

#### 2. 遗忘门：

遗忘门的作用是决定细胞状态中哪些信息应该被遗忘。它接收当前时刻的输入和上一时刻的隐藏状态作为输入，通过激活函数（通常是 Sigmoid 函数）输出一个介于 0 和 1 之间的数值。这个数值表示细胞状态中对应信息的保留概率，0 表示完全遗忘，1 表示完全保留。

#### 3. 输入门：

输入门用于决定当前输入的哪些信息应该被添加到细胞状态中。它同样接收当前时刻的输入和上一时刻的隐藏状态，通过 Sigmoid 函数输出一个控制信号，同时使用 tanh 函数生成一个候选值。控制信号与候选值相乘，得到要添加到细胞状态的信息。

#### 4. 输出门：

输出门根据当前的细胞状态和输入信息，决定最终的输出。它先通过 Sigmoid 函数生成一个控制信号，对细胞状态进行处理，再通过 tanh 函数将细胞状态映射到合适的输出范围，最后将两者相乘得到 LSTM 的输出。

以下是关于LSTM细胞状态 $C_t$ 、遗忘门 $f$ 、输入门 $u$ 、输出门 $o$ 的相关计算公式：

$$\begin{aligned}\tilde{c}^{<t>} &= \tanh(W_c[a^{<t-1>}, x^{<t>}] + b_c) \\ \Gamma_u &= \sigma(W_u[a^{<t-1>}, x^{<t>}] + b_u) \\ \Gamma_f &= \sigma(W_f[a^{<t-1>}, x^{<t>}] + b_f) \\ \Gamma_o &= \sigma(W_o[a^{<t-1>}, x^{<t>}] + b_o) \\ c^{<t>} &= \Gamma_u * \tilde{c}^{<t>} + \Gamma_f * c^{<t-1>} \\ a^{<t>} &= \Gamma_o * \tanh c^{<t>}\end{aligned}$$

Figure 1: LSTM的相关计算公式

LSTM的典型结构：

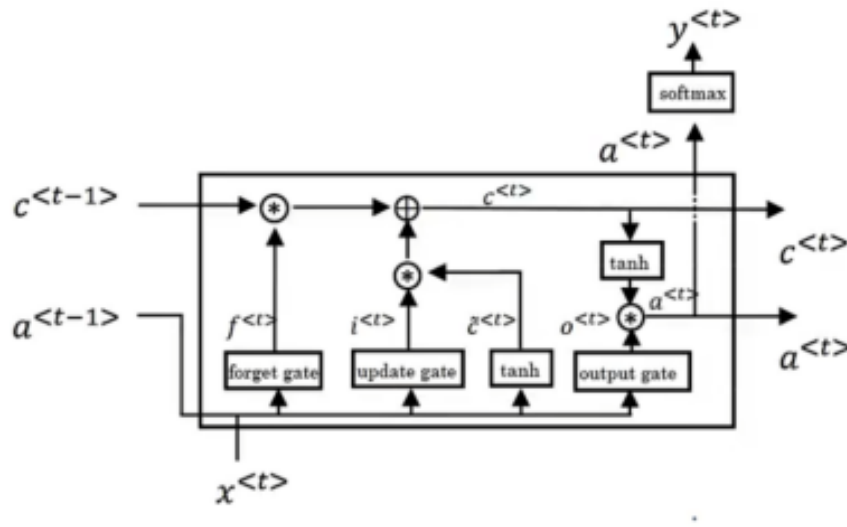


Figure 2: LSTM的典型结构

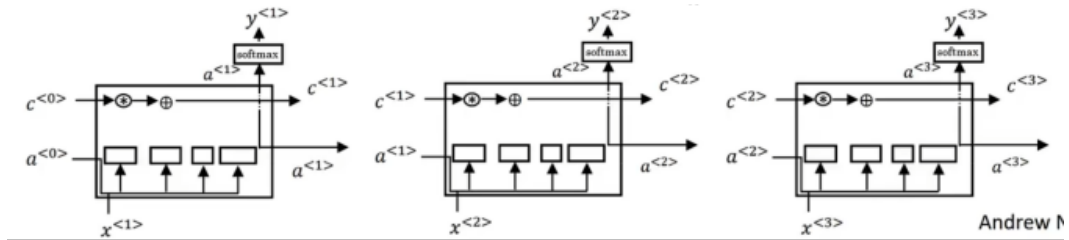


Figure 3: LSTM的典型结构

## 1.2 残差LSTM

残差LSTM是一种改进的循环神经网络结构，它通过将传统LSTM与残差连接相结合，有效解决了深层网络训练中的梯度消失和信息衰减问题。

残差LSTM保留了传统LSTM的三个核心门控机制（输入门、遗忘门和输出门），但在隐藏状态的计算中引入了残差连接。这种设计使得网络可以学习输入到输出的残差映射，而不是完整的变换，从而降低了学习难度。

### 1. 梯度消失:

在传统LSTM中, 随着网络深度增加, 反向传播时梯度需要经过多个时间步和非线性变换, 导致梯度逐渐衰减。残差连接通过建立“捷径”(shortcut path), 使梯度可以直接从深层回传到浅层, 有效缓解了梯度消失现象。

### 2. 网络退化:

当LSTM层数增加到一定程度时, 传统结构会出现性能饱和甚至下降的现象。残差连接通过让网络专注于学习输入输出之间的残差(即变化部分), 使深层网络更容易优化。

### 3. 信息传递中的特征衰减问题:

传统序列模型中, 原始输入特征经过多个时间步的变换后, 低层特征容易被高层特征淹没。残差连接的加法操作( $h_t = \text{传统输出} + \text{输入变换}$ )形成了一种特征叠加机制, 使得网络可以同时利用浅层局部特征和深层全局特征。

### 4. 长距离依赖建模不足:

对于需要捕捉远距离单词关系的任务(如情感分析中的“虽然...但是...”结构), 传统LSTM可能难以维持长期记忆。残差连接通过多路径传播机制, 将信息传递路径的最大长度缩短, 使模型对长距离的依赖关系识别更为准确。

因此, 为了解决这些问题, 残差LSTM应运而生。

残差LSTM在传统LSTM基础上增加了残差连接, 主要修改体现在隐藏状态计算, 门控信号计算与细胞状态更新同传统LSTM相同:

$$h_t = o_t \odot \tanh(C_t) + \mathcal{F}(x_t)$$

- $\mathcal{F}(x_t)$  是残差变换函数
- 当  $\dim(x_t) = \dim(h_t)$  时,  $\mathcal{F}(x_t) = x_t$  (恒等映射)
- 当维度不匹配时,  $\mathcal{F}(x_t) = W_{res}x_t$  (线性变换)

其中 $W_{res}$ 是可学习的权重矩阵, 用于调整输入 $x_t$ 的维度。

## 2 模型结构设计与损失函数选择

### 2.1 模型结构设计与参数设置

本实验采用了两种基于LSTM的文本情感分类模型: 基础LSTM模型和带有残差连接的LSTM模型(ResLSTM)。两种模型都遵循了相似的架构设计, 但在LSTM单元的实现上有所不同。

#### 1. 词嵌入层(Embedding Layer)

模型的第一层是词嵌入层, 将离散的单词索引转换为连续的向量表示。该层的输入维度为词汇表大小(vocab-size), 输出维度为嵌入维度(embed-dim=128)。使用padding-idx=0来指定填充符号的索引, 确保填充符号被映射为零向量。这一设计使得模型能够处理变长文本序列, 同时保留了单词的语义信息。

#### 2. LSTM层

基础LSTM模型使用标准的LSTM结构, 包含输入门、遗忘门和输出门三个门控机制。这种门控机制使LSTM能够有效地捕捉长距离依赖关系。

ResLSTM模型在标准LSTM的基础上引入了残差连接(residual connection)。具体实现是在计算最终隐藏状态时, 将当前时间步的输入 $x_t$ 经过线性变换(如果需要)后与标准LSTM的输出相加。这种设计使得梯度可以直接通过残差连接反向传播, 缓解了深层网络中的梯度消失问题。

#### 3. 情感分类层

LSTM层的输出经过Dropout层(丢弃率0.7)进行正则化后, 送入全连接层。全连接层将隐藏维度(hidden-dim=256)映射到单个输出值, 然后通过Sigmoid激活函数将输出压缩到(0,1)区间, 表示文本属于正面情感的概率。

## 2.2 损失函数选择

本实验选择了二元交叉熵损失(BCELoss)作为模型的损失函数，这是二分类任务的常用选择。其数学形式为：

$$L(y, \hat{y}) = -[y \log(\hat{y}) + (1 - y) \log(1 - \hat{y})]$$

其中 $y$ 是真实标签(0或1)， $\hat{y}$ 是模型预测的概率值。

### 1. 预测正确的情况

当真实标签为1且预测概率接近1时 ( $y=1, \hat{y} \rightarrow 1$ )，损失项为  $-\log(\hat{y})$  趋近于0，表示模型预测准确时惩罚极小。

当真实标签为0且预测概率接近0时 ( $y=0, \hat{y} \rightarrow 0$ )，损失项为  $-\log(1-\hat{y})$  同样趋近于0，反映正确预测负样本时的低惩罚特性。

### 2. 预测错误的情况

当真实标签为1但预测概率接近0时 ( $y=1, \hat{y} \rightarrow 0$ )，损失项  $-\log(\hat{y})$  会趋向无穷大，通过对数运算显著放大误差，产生强烈惩罚信号。

当真实标签为0但预测概率接近1时 ( $y=0, \hat{y} \rightarrow 1$ )，损失项  $-\log(1-\hat{y})$  同样急剧增大，确保模型能快速修正对负样本的错误判断。

因此，错误预测导致的损失增长幅度远大于正确预测的奖励幅度，并且预测与真实标签偏差越大，损失函数的梯度越显著，有利于模型快速修正严重错误。

## 3 实验结果与可视化

在相同训练环境下，LSTM最终准确率达81.45%，最佳准确率出现在第14个epoch；残差LSTM最终准确率为80.30%，准确率提升-1.15%，最佳准确率出现在第14个epoch。

```
===== 对比结果 =====  
LSTM最终准确率： 0.8145  
残差LSTM最终准确率： 0.8030  
准确率提升： -0.0115  
CNN最佳准确率出现在第 14 个epoch  
残差LSTM最佳准确率出现在第 14 个epoch
```

Figure 4: 结果对比

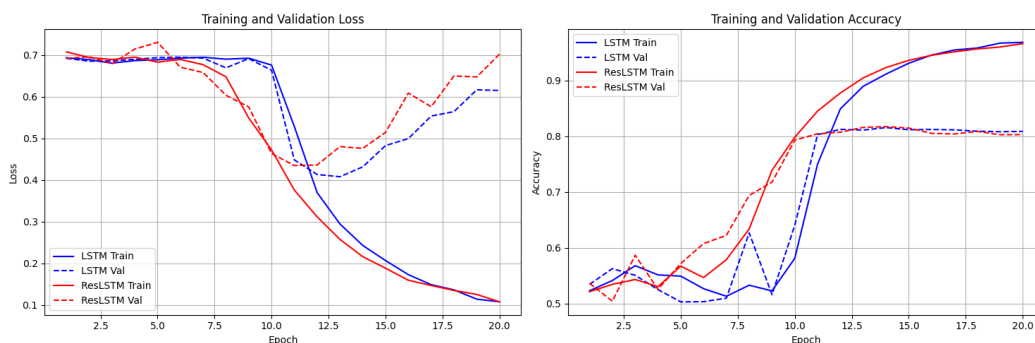


Figure 5: LSTM与残差LSTM可视化对比

## 4 总结与分析

实验结果表明，在IMDB情感分类任务中，基础LSTM模型（准确率81.45%）略优于残差LSTM模型（准确率80.30%）。这一现象可能与IMDB情感分类任务的特性相关：这类任务主要依赖于局部语义特征，如特定情感词汇的出现（如“great performance”或“terrible plot”），而不太需要处理复杂的长期依赖关系。在这种情况下，基础LSTM的门控机制已经足够捕捉关键特征，而残差连接引入的额外路径反而可能成为冗余计算。如图5所示，两种模型的训练曲线均呈现稳定收敛，但残差连接的加入并未带来预期提升。

通过图2和图3的LSTM结构可视化可见，标准LSTM的三门控设计（遗忘门、输入门、输出门）已能有效捕捉文本序列的时序特征。细胞状态的线性传播路径（图1公式中的 $C_t$ 计算）保证了关键信息的持久记忆，这可能是基础模型表现更优的结构基础。

虽然残差结构理论上能缓解梯度消失（如1.2节所述），但在本实验的浅层LSTM（单层结构）中，其优势未能充分显现。当网络深度不足时，残差连接可能引入冗余计算路径，反而轻微降低了参数效率。

本次实验为我提供了宝贵的经验：并非所有先进的网络结构改进都适用于所有任务场景。在实际应用中，需要根据具体任务特性选择合适的模型架构，盲目引入复杂结构可能会适得其反。