

# PRML 第三次实验报告（LSTM 空气污染预测）

22376367 黄正洋

## 1. 摘要 (Abstract)

本研究基于 LSTM (Long Short-Term Memory) 模型，面向 2010–2014 年北京市逐小时空气质量数据，提出一套用于未来一小时 PM2.5 浓度预测的多变量时间序列建模框架。模型采用滑动窗口机制（10 小时输入）与多维特征融合，结合 Min-Max 归一化、Dropout 正则化与早停机制，构建双层 LSTM 网络。实验表明，该模型在测试集上表现良好，取得  $RMSE = 0.07384$ 、 $MAE = 0.04752$ ，有效刻画出污染物随时间的变化趋势。

## 2. 引言 (Introduction)

PM2.5 浓度预测对公共健康和城市治理具有重要意义。传统预测方法（如 ARIMA、SVR）难以建模污染与气象之间的非线性关系。近年来，深度学习在时间序列领域表现突出，尤其是 LSTM 网络，因其引入门控机制缓解了梯度消失问题，能较好地捕捉长期依赖信息。本文基于 LSTM 网络构建污染预测模型，输入包括历史 PM2.5 浓度与温度、露点、风速等气象变量，目标是预测未来一小时的 PM2.5。

## 3. LSTM 模型结构与原理

### 3.1 网络结构

LSTM 网络引入记忆单元  $C_t$  与三个门控机制（遗忘门、输入门、输出门），其计算过程如下：

$$f_t = \sigma(W_f x_t + U_f h_{t-1} + b_f)$$

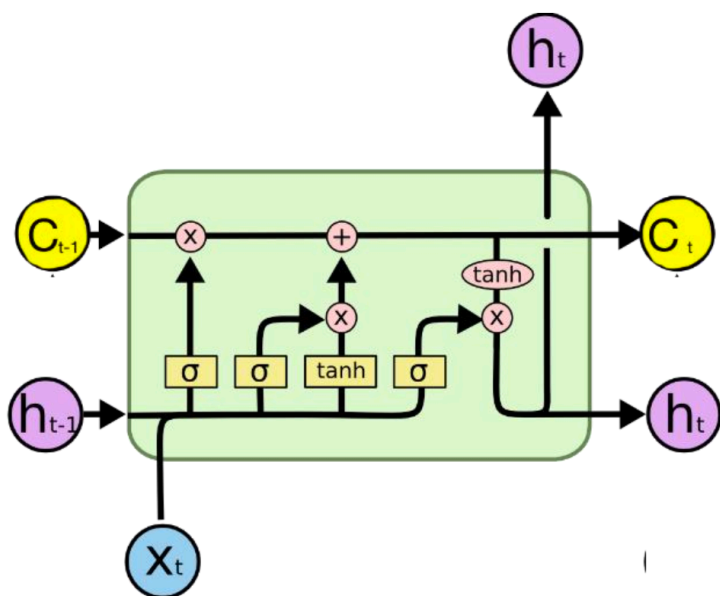
$$i_t = \sigma(W_i x_t + U_i h_{t-1} + b_i)$$

$$\tilde{C}_t = \tanh(W_c x_t + U_c h_{t-1} + b_c)$$

$$C_t = f_t \odot C_{t-1} + i_t \odot \tilde{C}_t$$

$$o_t = \sigma(W_o x_t + U_o h_{t-1} + b_o)$$

$$h_t = o_t \odot \tanh(C_t)$$



### 3.2 模型架构

层类型	参数配置	输出形状
LSTM-1	units=32, return_seq=True	(None, 10, 32)
Dropout-1	rate=0.2	(None, 10, 32)
LSTM-2	units=16, return_seq=False	(None, 16)
Dense	units=1	(None, 1)

## 4. 数据预处理

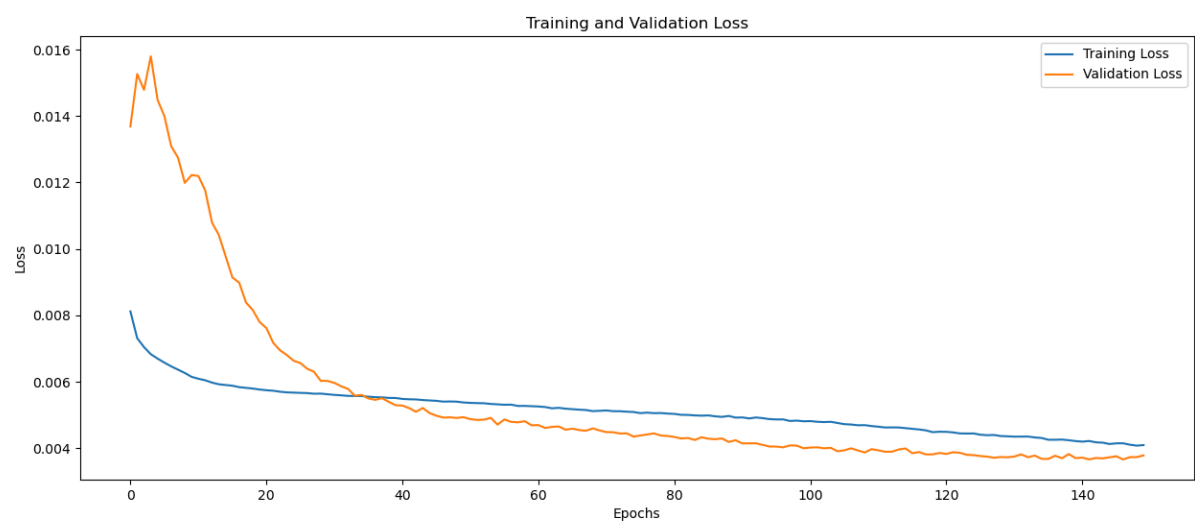
- 数据来源：Kaggle - Beijing PM2.5 Dataset，共 43800 小时数据（2010–2014 年）
- 字段包括：PM2.5、DEWP、TEMP、PRES、CBWD（风向）、lws、ls、lr
- 预处理操作：
  - 缺失值处理（删除 NA）
  - 风向 CBWD 映射为 0–3 整数
  - 滑动窗口生成输入  $(n - 10, 10, 7)$  和输出  $(n - 10, 1)$
  - Min-Max 归一化

## 5. 实验设置

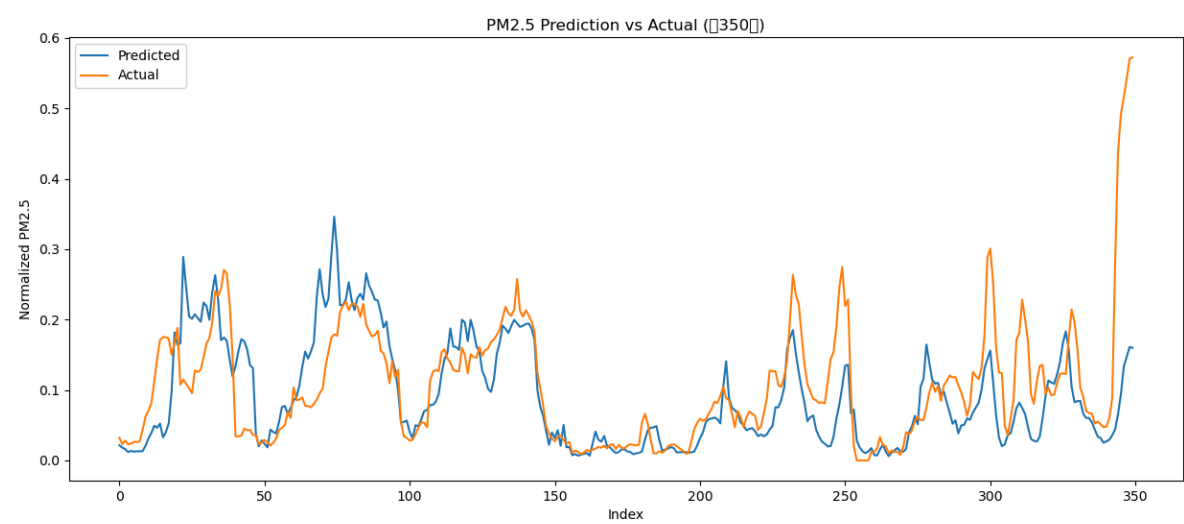
超参数	值
时间窗口	10 小时
Batch Size	32
Optimizer	Adam (学习率 0.001)
Epochs	最多 150，含早停机制
验证集	从训练集中划分 10%

## 6. 训练与结果可视化

### 6.1 损失收敛曲线



### 6.2 预测与实际值对比



### 6.3 指标评估

- RMSE: 0.07384
- MAE: 0.04752

模型预测曲线基本跟踪真实变化趋势，尤其在污染上升或下降区段拟合良好。

## 7. 模型评估与改进方向

局限	潜在优化方案
对突变污染响应滞后	引入注意力机制或 Transformer
模型泛化能力有限	增加外部变量如交通流量、时间戳编码
长期趋势建模效果不足	增加层数或尝试双向 LSTM

## 8. 总结

本文通过构建两层 LSTM 网络，实现对北京地区未来一小时 PM2.5 浓度的预测。模型融合多维气象特征并通过正则化与早停机制提升鲁棒性，在测试集上表现出较好精度（RMSE=0.07384，MAE=0.04752）。未来可从结构复杂度、上下文建模能力及数据增强方面进一步提升模型性能。

## 参考文献

1. Kaggle - Beijing PM2.5 Dataset: <https://www.kaggle.com/datasets/uciml/beijing-multivariate-time-series-data>
2. Hochreiter & Schmidhuber. Long Short-Term Memory. *Neural Computation*, 1997.
3. Colah's Blog: <https://colah.github.io/posts/2015-08-Understanding-LSTMs/>