

# DATA641 HW03 README

## 1. Python Environment

- Python version: 3.10 or higher
- Recommended platform: Google Colab (GPU enabled)

## 2. Requirements

- torch
- tensorflow
- scikit-learn
- pandas
- numpy
- matplotlib

## 3. How to Run

1. Open the notebook IMDB\_Sentiment.ipynb in Google Colab.
2. Upload the dataset IMDB\_Dataset.csv into /content/.
3. Run all cells sequentially — the notebook will:
  - a. Load and preprocess the IMDb dataset.
  - b. Build and train RNN/LSTM/BiLSTM models.
  - c. Run multiple experimental configurations.
  - d. Print metrics and generate comparison plots.
4. All output files will be saved under /content/results/.

## 4. Expected Runtime

Runtime Environment	Approx total time
Google Colab (CPU)	~45 - 60 minutes
Google Colab (GPU)	~12 - 18 minutes
Local GPU	~10 - 15 minutes

Each model (3 epochs) takes roughly:

- **LSTM / seq\_len=50:** ~45–50 seconds per epoch
- **BiLSTM / seq\_len=50:** ~85 seconds per epoch
- **RNN / seq\_len=50:** ~30 seconds per epoch

Here is the result when I used the Google Colab GPU:

Device: cuda  
CPU: x86\_64  
RAM (GB): 13.61  
CUDA: 12.6 | GPU: Tesla T4

... === Summary (last epoch per run) ===								
Model	Activation	Optimizer	Seq Length	Grad Clipping	Accuracy	F1	Epoch	Time (s)
BILSTM	RELU	ADAM	50	Yes	0.7510	0.7479		3.78
LSTM	RELU	ADAM	25	Yes	0.7033	0.7028	3.14	
LSTM	RELU	ADAM	50	No	0.7165	0.7164	2.71	
LSTM	RELU	ADAM	50	Yes	0.7208	0.7208	3.06	
LSTM	RELU	ADAM	50	Yes	0.7225	0.7177	3.06	
LSTM	RELU	ADAM	50	Yes	0.7315	0.7266	3.00	
LSTM	RELU	ADAM	50	Yes	0.6958	0.6803	3.08	
LSTM	RELU	ADAM	50	Yes	0.7479	0.7462	3.04	
LSTM	RELU	RMSPROP	50	Yes	0.5001	0.3349	3.08	
LSTM	RELU	SGD	50	Yes	0.5021	0.4696	2.75	
LSTM	SIGMOID	ADAM	50	Yes	0.7071	0.7069	3.11	
LSTM	TANH	ADAM	50	Yes	0.7515	0.7515	3.00	
LSTM	RELU	ADAM	100	Yes	0.7481	0.7415	3.19	
RNN	RELU	ADAM	50	Yes	0.5608	0.4819	3.14	
== Best (by F1) ==								
Model	Activation	Optimizer	Seq Length	Grad Clipping	Accuracy	F1	Epoch	Time (s)
LSTM	TANH	ADAM	50	Yes	0.7515	0.7515	3.0	
== Worst (by F1) ==								
Model	Activation	Optimizer	Seq Length	Grad Clipping	Accuracy	F1	Epoch	Time (s)
LSTM	RELU	RMSPROP	50	Yes	0.5001	0.3349		3.08

Just changing the Runtime type in Colab to GPU, there is not much change in trends. The Report and code I submitted contains the results and comparisons while running on the CPU.

