



Vidyavardhini's College of Engineering and Technology
Department of Artificial Intelligence & Data Science

AY: 2025-26

| | | | |
|--------------|----|--------------|-----|
| Class: | AI | Semester: | VII |
| Course Code: | | Course Name: | DL |

| | |
|----------------------|------------------|
| Name of Student: | BARI ANKIT VINOD |
| Roll No. : | 61 |
| Assignment No.: | 5 |
| Title of Assignment: | |
| Date of Submission: | |
| Date of Correction: | |

Evaluation

| Performance Indicator | Max. Marks | Marks Obtained |
|------------------------|------------|----------------|
| Completeness | 5 | |
| Demonstrated Knowledge | 3 | |
| Legibility | 2 | |
| Total | 10 | |

| Performance Indicator | Exceed Expectations (EE) | Meet Expectations (ME) | Below Expectations (BE) |
|------------------------|--------------------------|------------------------|-------------------------|
| Completeness | 5 | 3-4 | 1-2 |
| Demonstrated Knowledge | 3 | 2 | 1 |
| Legibility | 2 | 1 | 0 |

Checked By

Name of Faculty : Raunak Joshi

Signature :

Date :

Assignment No. - 5

DL

Q. i) Discuss the way LSTM solve the vanishing gradient problem of RNN.

→ (i) The vanishing gradient problem in RNN's -

- standard RNN's process sequences by recursively applying the hidden problem -

$$h_t = \tanh(W_h h_{t-1} + W_x x_t + b)$$

- during backpropagation through time (BPTT), the gradient involves state:

$$\frac{\partial L}{\partial h_t} = \frac{\partial L}{\partial h_t} \prod_{k=t}^{T-1} \frac{\partial h_{k+1}}{\partial h_k}$$

- problem: Since $|\tanh(x)| < 1$, multiplying many small numbers.

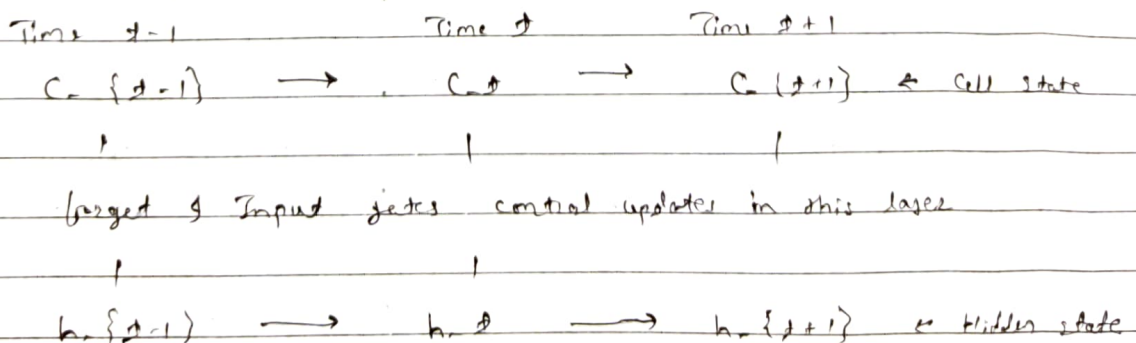
(ii) How LSTM solves it -

- LSTM (Long-Short-Term-Memory) introduces a special cell state C_t and gates that control introduction flow.

key components -

- Cell state C_t - Acts as a highway for gradients.
 - gradient can unchanged over many time steps.
- Gates :- forget gate f_t
 - input gate i_t
 - output gate o_t

(iii) Visual Representation -



∴ The cell state acts like a "gradient highway", letting long-term information pass without decay.

Q. 2) Compare and contrast RNN and LSTM for Text Classification Task.

| RNN | LSTM |
|---|--|
| Simple recurrent unit with hidden state h_t | LSTM cell with hidden h_t and cell state C_t |
| Only short term memory | forget, input, output, gates. |
| Poor - suffers from vanishing / exploding gradients | excellent - gates + additive cell state preserve gradients. |
| may fail when context spans long sequences. | can retain important contextual information across long sequences. |
| less stable, may require gradient clipping. | more stable, better convergent on long sequences. |
| dropout, or recurrent dropout required. | Can use some plus intrinsic gate regularization. |
| works well if sequences are short. | works well too; overkill for every short text. |
| Poor performance on long sequences. | superior performance, retains context from earlier words. |
| limited context \rightarrow less accurate | rich contextual representation \rightarrow better accuracy. |
| faster, simpler | slower, heavier due to gates. |