Iran University of Science & Technology

School of Computer Engineering

Assignment #6

Neural Networks

DR. Marzieh Davoodabadi, Fall 2024

Teaching Assistants:

MohammadAmin Qolizade

Due: 1403/10/21

Contents

Notes:

1. Submit the answers in a complete PDF file and the code for the questions in the .ipynb format (including the notebook cell outputs) in a compressed file named HW6_StudentID.zip by the specified deadline.
2. A total of 72 + 72 hours of delay in submitting the answers is allowed across all projects. After that, for each additional day of delay, 10% of the score will be deducted.
3. If a student submits the project earlier than the deadline and achieves 75% of the score, up to 24 hours will be added to their allowable delay time.
4. The maximum delay for submitting each assignment is 5 days, and after 5 days, submission will not be accepted.
5. It is important to note that the explanation of the code and the obtained results must be included in the PDF file. Code without a report will result in a score deduction.
6. The evaluation of the assignment will be based on the correctness of the solution and the completeness and accuracy of the report.
7. Assignments must be completed individually, and group work on assignments is not allowed.
8. Please allocate sufficient time for the assignment and avoid leaving it until the last days.
9. You can ask your questions in the relevant group.

good luck.

Problem 1

Consider a simple RNN with the following parameters: **(15 points)**

- Input at time step t = 1: $x_1 = [0.5, -0.2]$
- Weight matrices:

$$w_h = \begin{bmatrix} 0.4 & 0.6 \\ -0.2 & 0.1 \end{bmatrix}, \quad w_x = \begin{bmatrix} 0.3 & 0.1 \\ 0.7 & -0.3 \end{bmatrix}$$

- Initial hidden state $h_0 = [0, 0]$

Use the formula for the hidden state update in an RNN:

$$h_t = \tanh(W_{xh}x_t + W_{hh}h_{t-1})$$

a. Calculate the hidden state $h_1 = y_1$ at time step t = 1.
b. Explain the working principle. Describe the process of forward propagation in RNN, focusing on how the hidden state is updated at each time step.

Problem 2

Consider an LSTM with the following parameters at time step t = 2: **(30 point)**

- Previous hidden state $h_1 = [0.3, 0.6]$
- Previous cell state $C_1 = [0.5, -0.3]$
- Input at time step t = 2: $x_2 = [-0.1, 0.4]$
- Weight matrices for the gates:

$$w_f = \begin{bmatrix} 0.2 & -0.4 \\ 0.1 & 0.5 \end{bmatrix}, w_i = \begin{bmatrix} -0.3 & 0.2 \\ -0.4 & 0.3 \end{bmatrix}, w_c = \begin{bmatrix} 0.6 & -0.1 \\ 0.5 & 0.3 \end{bmatrix}, w_o = \begin{bmatrix} -0.2 & 0.1 \\ 0.3 & -0.5 \end{bmatrix}$$

- Bias vectors for the gates:

$$b_f = [0.1, -0.2], \quad b_i = [-0.1, 0.2], \quad b_c = [0.3, -0.1], \quad b_o = [0.05, -0.05]$$

Use the following formulas to update the LSTM cell:

- Forget gate:

$$f_t = \sigma\left(W_f \begin{bmatrix} h_{t-1} \\ x_t \end{bmatrix} + b_f\right)$$

- Input gate:

$$g_t = \tanh\left(W_c \begin{bmatrix} h_{t-1} \\ x_t \end{bmatrix} + b_c\right) \qquad i_t = \sigma\left(W_i \begin{bmatrix} h_{t-1} \\ x_t \end{bmatrix} + b_i\right)$$

- Cell state update:

$$C_t = i_t \circ g_t + f_t \circ C_{t-1}$$

- Output gate:

$$o_t = \sigma\left(W_o \begin{bmatrix} h_{t-1} \\ x_t \end{bmatrix} + b_o\right) \qquad h_t = o_t \circ \tanh(C_t)$$

a. Calculate the forget gate $\mathbf{f}_2$, input gate $\mathbf{i}_2$, candidate cell state $\mathbf{g}_2$, and the cell state $\mathbf{C}_2$.
b. Calculate the output gate $\mathbf{o}_2$ and the hidden state $\mathbf{h}_2$ at time step t = 2.
c. Describe the structure of an LSTM unit and explain the role of each component: the input gate, forget gate, and output gate. How do these gates help address the limitations of basic RNNs?

Problem 3

The notebook file HW6_Q3 includes the implementation of a weather prediction model using LSTM with Keras. **(20 point)**

a) Study the code to understand its functionality and how the model works. Write a report detailing your analysis and understanding of the code.
b) Implement the model using GRU and SimpleRNN, and compare the results obtained from each approach.

Problem 4

Based on what you have learned so far, use the dataset file "text.csv" and the notebook HW6_Q4 to complete the tasks requested in the notebook (TODO). Please note that all topics outside the scope of the course have already been implemented, and your task is to complete the sections required to finalize the "train" part. **(35 point)**

The report for this section should include a detailed analysis of the data, the code's outputs, and a full explanation of the completed code sections.