Shiraz University

**Neural Networks and Deep Learning** (Ordibehesht 1402)

Assignment #2: Recurrent Neural Network

**Due date: 31th Ordibehesht 1402**

In this project, you are going to get familiar with single-hidden layer recurrent networks.

You will design a text emotion detection system by using the Elman network. You have a dataset containing 3250 text rows with 5 classes. Split last 150 text of each class for the test dataset.

1. If there are any non-letter characters in each row of the text column, remove them.
2. Then, tokenize them by converting each text to word sequences.
3. Remove short words (with length $\leq 2$).
4. Remove all stop words (e.g., 'a', 'and', 'what', …), given in the file 'stopwords.txt'.
5. Add padding to ensure that all the sequences have the same length (taking max length).
6. By using the one-hot encoding method, convert the words of each sequence into a numerical vector. (In this phase, you need to make a dictionary with unique words from all texts.)
7. Build a many-to-one RNN model with Elman Network.
8. Train your model with your train dataset and report the accuracy of your model on train and test data.

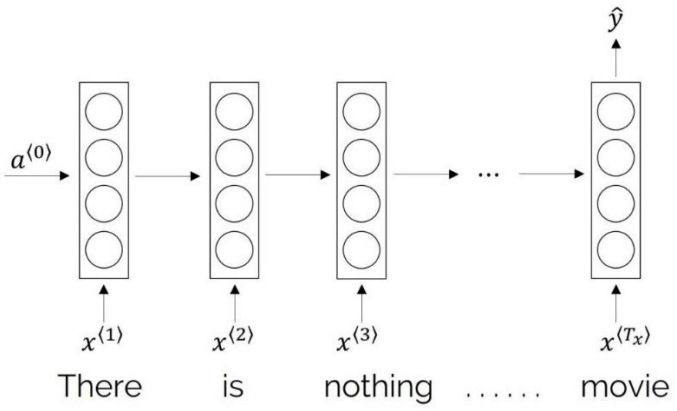


*Figure 1: Example of Many-to-One architecture for text analysis*

How this problem can be solved using Hopfield recurrent network. Explain your idea.

Notes:

- Pay extra attention to the due date. It will not extend.
- Be advised that submissions after the deadline would not grade.
- Prepare your full report in PDF format and include the figures and results.
- Feel free to use any predefined functions.
- Email your folder in this format (HW#_student#_name_family.zip) to **soroushmehrpou@gmail.com**