Indian Institute of Technology Hyderabad
Deep Learning (AI2100/AI5100): Assignment-4

Topic: RNNs
Assigned on: 26^{th} **March, 2023**
Deadline: 10^{th} **April, 2023**
Maximum Marks: 40

1 Instructions

- Answer all questions. We encourage best coding practices by not penalizing (i.e. you may not get full marks if you make it difficult for us to understand. Hence, use intuitive names for the variables, and comment your code liberally. You may use the text cells in the notebook for briefly explaining the objective of a code cell.)
- It is **expected** that you work on these problems individually. If you have any doubts please contact the TA or the instructor no later than 2 days prior to the deadline.
- You may use built-in implementations only for the basic functions such as `sqrt, log`, etc. from libraries such as `numpy` or `PyTorch`. Other high-level functionalities are expected to be implemented by the students. (Individual problem statements will make this clear.)
- For plots, you may use `matplotlib` and generate clear plots that are complete and easy to understand.
- You are expected to submit the Python Notebooks saved as <your-roll-number>.ipynb
- If you are asked to report your observations, use the mark down text cells in the notebook.

2 Questions

1. **The adding problem**: In this task, each data sample consists of a sequence of variable length, but a constant depth of 2. All values of the first dimension (randomly) lie in $[0, 1]$, and the second dimension being all zeros except for two elements that are marked by 1. Objective of the task is to sum the random values whose second dimensions are marked by 1. Train the different RNNs (Elmon network, LSTM, and GRU) discussed in the class and compare their performance against a baseline that always predicts a sum of 1 plotting the learning curves and final performance. Note that you are expected to implement these models (as opposed to use the builtin constructs).[7 (Elmon)+12 (LSTM)+ 12(GRU)+5(baseline&comparison) = 36]
   The following table presents two data samples (x) along with their labels (y). Note that the samples should be of different lengths (n), so the dimensions of each sample can be represented as $n \times 2$. Given examples have lengths of 5 and 8 respectively. You have to generate/create a big dataset (≥ 5000) of such samples for training and testing of the RNNs. [Dataset creation 4 Marks]

	x								y
1	0.1 0	0.9 1	0.25 0	0.17 1	0.76 0				1.07
2	0.86 0	0.31 0	0.43 1	0.12 0	0.01 0	0.29 0	0.95 0	0.09 1	0.52