Indian Academy of Sciences, Bengaluru
Indian National Science Academy, New Delhi
The National Academy of Sciences India, Prayagraj

SUMMER RESEARCH FELLOWSHIPS — 2021

Format for the Four-week Report*,^

Name of the candidate	:	Kushagra Bansal
Application Registration no.	:	MATS398
Date of Commencement of work	:	07/06/2021
Mode of work	:	From Home: ✓ ☐ Guide's Laboratory: ☐
Name of the guide	:	Prof. Shayan Srinivasa Garani
Guide's institution	:	Indian Institute of Science, Bengaluru
Place of stay during the tenure of the fellowship (if working in guide's institution)	:	Hostel provided by Guide: - Own arrangement: - Other (Specify): -

kushagra

Signature of the candidate

Date: 06/07/2021

Shayan GS

Signature of the guide

Date: 26/07/2021

INSPIRE/KVPY FELLOWSHIP (please fill this box)#			
1.	I am currently a recipient of	INSPIRE FELLOWSHIP	☐ Yes / ☑ No
		KVPY FELLOWSHIP	☐ Yes / ☑ No
		If, YES, fill cols. 2, 3 & 4	
2.	**INSPIRE/KVPY** Fellowship is from ___-___[month]/___-___[yr] to ___-___[month]/___-___[yr]		
3.	I receive a monthly fellowship of Rs. ____-_____ from INSPIRE/KVPY towards my living expenses		
4.	I also receive towards contingencies a sum of Rs. _____-_____ per year		
	I affirm that the information given above is correct. kushagra **Signature of the candidate**		

(For office use only; do not fill/tear)

Candidate's name:		Fellowship amount:
Student:	Teacher:	Deduction:
Guide's name:		Amount to be paid:
KVPY Fellow:	INSPIRE Fellow:	A/c holder's name:
PFMS Unique Code:		Others:

IMPORTANT NOTES:

*** The four-week report could be between 300 and 350 words.**

^ This format should be the first page of the report and should be stapled with the main report.

 # Mandatory to fill this section, this should be filled and signed by you even if you are not an INSPIRE/KVPY Fellow. Otherwise release of fellowship amount will be withheld.

Four Week Progress Report

Kushagra Bansal

July 06, 2021

As part of my fellowship, I am currently working under Prof. Shayan Srinivasa Garani of the Indian Institute of Science, Bengaluru. My topic for research is recurrent neural networks (RNN) and long short term memory networks (LSTM). I initially started with Alex Sherstinsky's paper "Fundamentals of Recurrent Neural Network and Long Short Term Memory Networks". The paper derives the canonical equations for an RNN cell by using the additive model from brain dynamics and applies backward Euler propagation to arrive at the equations.
After deriving the equations, the paper proposes a proposition which yields the unrolling technique for a recurrent neural network model.
Moving on, we then look at the training equations for an RNN by using the technique called "backpropagation through time" (BPTT). This leads us to investigate the well-known training difficulties known as "vanishing and exploding gradients".
To overcome these issues while training a recurrent neural network, we supply some logical arguments to construct a modified RNN cell which gradually evolves into what is known as "vanilla long short-term memory" network.
In the following section, we look at the governing equations of the Vanilla LSTM network in detail by highlighting all the parameters of the model.
Just like for Recurrent neural network models, we then look at the training side for a LSTM network. The technique "backpropagation through time" is employed here. We define some intermediate variables to arrive at the equations.
After that we mathematically prove how do the modification in the RNN cell, which leads to an LSTM cell, circumnavigate the vanishing gradient problem.
Then in the last section we look at a few modifications of the "vanilla LSTM

model" such as the "augmented LSTM system". The author concludes the paper by stating further research prospects in the vast topic.
After carefully studying this paper and a few more resources, I was given the task to develop an LSTM network for music generation. For developing that, I first stared with implementing a simple RNN model for text generation. My RNN model has 100 hidden layer nodes. I have used the cross-entropy loss function to evaluate performance. The training of the model is done by BPTT algorithm and Adagrad update is used to update the model parameters. It is developed in Python3 using on the Numpy library to easily handle vectors and matrices. The complete details for this is put in a separate write up.
For the next four weeks, my focus shall be on evolving this simple RNN model into an LSTM model adapted for music generation. I also wish to experiment with different loss functions and update rules to document and analyse the different results produced by them.