

# Assignment 3

CS5370: Deep Learning for Vision  
IIT-Hyderabad  
Jan-Apr 2017

**Max Marks:** 40  
**Due:** 19th Apr 2017 11:59 pm

## Instructions

- Please use Google Classroom to upload your submission by the deadline mentioned above. Your submission should comprise of a single ZIP file, named `<Your_Roll_No>Assign3`, with all your solutions, including code.
- For late submissions, 10% is deducted for each day (including weekend) late after an assignment is due. Note that each student begins the course with 6 grace days for late submission of assignments. Late submissions will automatically use your grace days balance, if you have any left. You can see your balance on the CS5370 Marks and Grace Days document under the course Google drive.
- Please read the department plagiarism policy. Do not engage in any form of cheating - strict penalties will be imposed for both givers and takers. Please talk to instructor or TA if you have concerns.

## 1 Programming (40 marks)

The code for the assignment is uploaded alongside this document on the Classroom portal (`assign3code.zip`)<sup>1</sup>. An RNN is well suited for tasks on

---

<sup>1</sup>Thanks to Dr Ankit Patel at Rice University for this code and assignment design

time-series or sequential data. However, in this assignment, we will see how we can use RNN for object recognition.

1. With a CNN, we would send in full images; but with this RNN, we will input 28 pixel chunks of the images to the network. For example, each row of the image will be sent to the network at each iteration. The starter code, `lstmMNISTStarterCode.py`, provided will detail how it splits the data so that it can work for the RNN. You should modify the following parameters in the starter code, and see how it works:
  - Number of nodes in the hidden layer
  - Learning rate
  - Number of iterations
  - Cost (*Hint*: use softmax cross entropy with logits)
  - Optimizer

Write a short report on your observations.

2. Change line 35 in the starter code (see below) to use LSTM and GRU instead of RNN.  
`rnn_cell.BasicRNNCell(nHidden)`  
Plot the train/test accuracies and the train loss. What do you notice - are the LSTM or GRU better? Also, change the number of hidden units and see how that affects the loss and accuracy.
3. Compare with training using a CNN instead for the same task and describe any similarities or differences.

## 2 Practice Exercises

**NO SUBMISSION REQUIRED; PLEASE USE THIS FOR PRACTICE AND LEARNING.**

1. **Image captioning:** Implement Assignment 3 (<http://cs231n.github.io/assignments2016/assignment3/>) on the Stanford CSC231n course. (We are not doing this mandatorily due to the time left in the course, and the need for GPUs in this assignment).
- 2.