# Indian Institute of Technology Hyderabad
# Deep Learning (AI2100/AI5100): Assignment-3

**Topic**: Gradient Descent Update rules and CNN Visualization
**Assigned on:** $2^{nd}$ **March, 2023**
**Deadline:** $16^{th}$ **March, 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. **Variants of Backprop for visualizing CNNs**
   Download a ResNet-50 trained on the ILSVRC classification task and 5 images from here. Considering the neuron corresponding to the highest predicted score in the output layer, implement and visualize the results for
   (a) Vanilla Backpropagation (you may use the builtin implementation) (2)
   (b) Guided-back-propagation (3) (you may use the builtin grad function for computing gradient of a node with other in the computational graph)
   (c) Class Activatoin Map (CAM) (4)
   (d) Guided Grad-CAM (6)
   Clear visualization of these results for the chosen 5 images carries 5 marks.

2. **SGD Update rule variations**
   Implement the following update rules to find the minimum for $f(x, y) = 10x^2 + y^2$. You may consider a random starting point and visualize the convergence (within the region covered by points (-2, -2) and (2, 2) using a GIF) path taken by different update rules for a fixed (say, 50) iterations. You are advised to tryout multiple meaning values for the hyper-parameters involved.
   (a) Vanilla SGD (2)
   (b) SGD with momentum (3)
   (c) Nesterov Momentum (3)
   (d) AdaGrad (3)
   (e) RMSProp (4)
   Clear convergence visualization using a GIF image carries 5 marks.