**Assignment2\_Data\_Instructions**

1. **Task 1 :** Use the assignment 1 datasets (Linearly Separable, Non-linearly Separable, and Overlapping Data ) as per your group.
2. **Task 2, 3, and 4:** : Data is hosted on [10.6.8.2/~naresha/kmpa/](http://10.6.8.2/~naresha/kmpa/)Task2\_3\_4-data
   1. Download your group folder only. [Use **wget -r <pathtoyourgroup>**]
   2. Load  **‘Group\_<no>\_data.mat’**
   3. **labelName** variable contains the class names assigned to your group.

(Each group is assigned with 5 classes)

* 1. **final\_labels** has the following information**:**  the first column indicated the **image no.** and the remaining 5 columns represents the class membership of in **1-of-5 representation.**
  2. You can see the images in your group folder.
  3. **Task2 Feature Extraction Part**: Run the script file (**featureextraction\_task2.p**) to get the ‘**feature representation**’ for images.

Also include the **LMgist.p** and **imresizecrop.p**  file to execute the script file.

Usage of script file:

Function takes image as input and returns the features as output.

Eg: [feature\_ouput]=**featureextraction\_task2(‘1.jpg’);**

g. **Task3 and Task4:** The original images are size of 200\*200.

Reduce the size of images to 32\*32 before you train CNN.

1. **Task 5:**
   1. Load the **binary\_images.mat** file
   2. **data** variable contains binary features.
   3. Take your group data based on the classes assigned to you. Group mapping file can be seen in folder **Task5\_data**)
   4. Take examples corresponding to your classes.(use simple indexing to get your data). Examples to classes mapping is available in **labels** variable.

Split the data into training, validation and test with 60:20:20 (preferably!) wherever it’s required.

**Suggested Toolbox:**

**Task 1:** Matlab + LibSVM (software is available in gdrive) [Installation instructions for LibSVM: <http://www.cs.cmu.edu/~guestrin/Class/15781/hws/hw3/libsvm_inst.txt>]

**Task 2, 3, 4, 5**: Torch (Manual: <https://github.com/torch/torch7/wiki/Cheatsheet>)

**Note: Any issues with data sets, please write to** [*kernel.jan.july.2016@gmail.com*](mailto:kernel.jan.july.2016@gmail.com) or contact Naresh Annam.

Start assignment early, any difficulty get clarified as early as possible! ATB.