* Better understanding of the code:
  + main files: DeepDPM, clusternetasmodel, classifiers
  + main args (gpus, use\_labels\_for\_eval)
  + log
* Go deep – **Clustering net**:
  + Map Supplemental Material (chapter 7):
    - input layer, a single hidden layer, and an output layer
    - The number of hidden units was always 50 in all our experiments (changing that number had little effect on the results).
    - The (changing) number of neurons in the output layer was K
* Go deep - **Subclustring net**:
  + SHAPE is input dim(from clustering net) -> Hidden\_dim \* K -> 2 \* K
  + 2 FC layes
  + Hidden\_dim = 50 is hard coded
  + detach different subclustering nets - zeroing out the weights connecting between different subnets. And also zero their gradient
  + Basic changes I done:
    - 1- Run:
      * Mnist, Initial K = 1
      * Hidden\_dim = 50
      * Add dropout layers between the Fc layers with drop = 0.5
      * ->Subclustring Params: 652
      * Result: NMI: 0.94161, ARI: 0.95366, acc: 0.9787, final K: 10
    - 2- Run:
      * Mnist, Initial K = 1
      * Hidden\_dim = 75
      * Add dropout layers between the Fc layers with drop = 0.5
      * ->Subclustring Params:977 (13 params for each additional neuron?)
      * Result: NMI: 0.94165, ARI: 0.95372, acc: 0.97873, final K: 10
    - 3- Run:
      * Fashion imbalanced, Initial K = 1
      * Hidden\_dim = 75
      * Add dropout layers between the Fc layers with drop = 0.5
      * ->Subclustring Params:977
      * NMI: 0.64168, ARI: 0.49487, acc: 0.60146, final K: 10
    - 4- Run:
      * Fashion imbalanced, Initial K = 1
      * Hidden\_dim = 75
      * Add dropout layers between the Fc layers with drop = 0.2
      * ->Subclustring Params:977
      * NMI: 0.6437, ARI: 0.49242, acc: 0.5972, final K: 9
    - 5- Run:
      * EPOCHS: 700
      * Fashion imbalanced, Initial K = 1
      * Hidden\_dim = 75
      * Add dropout layers between the Fc layers with drop = 0.2
      * ->Subclustring Params:977
      * NMI: 0.64513, ARI: 0.4967, acc: 0.61412, final K: 9
  + Advanced Changes:
    - **Adding one more layer of hidden dim neurons (updating merge and split according to that)**
    - 1 – Run:
      * Mnist, Initial K = 1
      * Epochs = 600
      * Hidden\_dim = 50
      * Drop = 0
      * >Subclustring Params: 3200
      * NMI: 0.86102, ARI: 0.69085, acc: 0.78364, final K: 8
    - 2 – Run:
      * Mnist, Initial K = 1
      * Epochs = 500
      * Hidden\_dim = 25
      * Drop = 0
      * >Subclustring Params: 3200
      * NMI: 0.94162, ARI: 0.95369, acc: 0.97871, final K: 10
    - 3 – Run:
      * Mnist, Initial K = 1
      * Epochs = 500
      * Hidden\_dim = 25
      * Drop = 0.3 (between dc1,fc2)
      * >Subclustring Params: 977
      * NMI: 0.94161, ARI: 0.95366, acc: 0.9787, final K: 10
    - 4 – Run:
      * Mnist\_IMBALANCED, Initial K = 1
      * Epochs = 500
      * Hidden\_dim = 25
      * Drop = 0.3 (between dc1,fc2)
      * >Subclustring Params: 977
      * NMI: 0.95029, ARI: 0.97132, acc: 0.98237, final K: 10
    - 5 – Run:
      * Mnist\_IMBALANCED, Initial K = 1
      * Epochs = 500
      * Hidden\_dim = 15
      * Drop = 0.4 (between dc1,fc2)
      * >Subclustring Params:437
      * NMI: 0.95024, ARI: 0.97113, acc: 0.98245, final K: 10
    - 6 – Run:
      * Fashion\_IMBALANCED, Initial K = 1
      * Epochs = 500
      * Hidden\_dim = 15
      * Drop = 0.4 (between dc1,fc2)
      * >Subclustring Params:437
      * NMI: 0.64681, ARI: 0.49417, acc: 0.60761, final K: 8
    - 7– Run:
      * Fashion\_IMBALANCED, Initial K = 1
      * Epochs = 700
      * Hidden\_dim = 25
      * Drop = 0.4 (between dc1,fc2)
      * >Subclustring Params: 977
      * 0.64645, ARI: 0.49401, acc: 0.60795, final K: 8
    - 8 – Run:
      * Mnist, Initial K = 1
      * Epochs = 500
      * Hidden\_dim = 15
      * Drop = 0
      * >Subclustring Params: 437
      * NMI: 0.94161, ARI: 0.95366, acc: 0.9787, final K: 10