M2015: Machine Learning: Assignment 3
Introduction to Deep Learning

IIIT-Hyderabad
Instructor:Prof.C.V.Jawahar

General Instructions

- This assignment is going to be a little time consuming and you might have to set-aside your laptop for 5-6 hours to learn the model. So plan and start the assignment early.
- The tutorial would cover the basics of what needs to be done for Section 2 and Section 3
- Limit your answers for Section 1 and Section 4 to 2 A4 size pages (typed) each.

1 Discussion on one recent success of deep learning

Deep learning has been constantly causing shock-waves in the field of machine learning, achieving and surpassing the performance of most traditional models. It finds its applications in a wide variety of applications from speech to natural language processsing, from robotics to computer vision. In this part of the assignment we expect you to explore the different applications on which deep learning has had a significant impact. The goal is as follows:

Discuss a recent success of deep learning in any of the application areas, such as, but not limited to, computer vision, NLP, speech processing, etc. You need to mention the following points:

1. What are the novelties introduced in this paper?
2. Illustrate and explain the architecture of the network used?
3. What are the assumptions made regarding the problem to be solved?
4. Why do you think the method beats the baseline methods in the paper?
5. What are the possible drawbacks? Can you suggest suitable modifications?
6. Can you think of an extension to this application or some other application which can use a similar architecture?

2 Convolution Neural Network using CIFAR-10 dataset

Convolution neural networks(CNNs) are a variant of the traditional multi-layer perceptron. In this part of the assignment you will learn the following:

- Independently train a convolution neural network.
- Observe and compare the power of the representations learnt using this network with that of raw pixels

We would use the CIFAR-10 [1] dataset for this section. The CIFAR-10 dataset consists of 60000 32x32 colour images in 10 classes, with 6000 images per class. There are 50000 training images and 10000 test images.

Run the following experiments:

e_1 Train a CNN using raw pixels as features with a softmax classifier on top for 100 epochs

e_2 Using the trained model from e_2, chop off the softmax classifier. Extract features from this model, and train a SVM using these features.

After computing the confusion matrix for both train and test sets for each of e_1 and e_2, report the following:

- Overall accuracy of the classifier
- Label-wise accuracy of the classifier
- Does training a SVM using features from a pre-trained CNN help?
- Extract test features from the trained model after every 10 epochs, and use tSNE to visualise the classes. Show how the separability of the classes improves as training progresses.

3 Parameter tuning using CIFAR-100 dataset

In this part of the assignment you would learn how to adapt the representations learnt on a dataset to another "similar" dataset. The goals are as follows:

- Adapt the representations learnt to perform on a similar dataset by fine tuning the parameters.
- Realise how easy/difficult it is to tune the parameters of the network to achieve the desired optima.

We would use the CIFAR-100 [1] dataset. This dataset is just like the CIFAR-10, except it has 20 classes. Thus, the hope is that the representations learnt using CIFAR-10 would be rich enough to distinguish the 20 classes. We use the representation learnt using the CIFAR-10 dataset and apply a 20 class classifier on top of it, say softmax classifier.

You need to perform the following experiments:

e_3 Using the trained model from e_1, chop off the softmax classifier and the last fully connected layer. Add a new fully connected layer, along with a new softmax classifier, and retrain the model on the new data for about 25 epochs. Repeat this experiment at least thrice with different learning rates for the new layers. You can add multiple new layers if you want.

[1] //www.cs.toronto.edu/ kriz/cifar.html

e_4 Train a CNN using raw pixels as features with a softmax classifier on top for 25 epochs only.

After computing the confusion matrix for both train and test sets for each of e_3 and e_4, report the following:

- Overall accuracy of the classifier
- Show error and objective plots for both models
- Discuss how and why finetuning performs better or worse than training from scratch in this scenario.

4 Advances made in Convolution Neural Networks and their performance on the CIFAR-10 and CIFAR-100 dataset

The convolution neural networks which we used in the previous stage is a vanilla version. However, various modifications have been done on this network. A large number of experiments have been done by tweaking the number of layers, the size of convolution filter, the strides it take, pooling etc. The goal of this section is to read and summarise one such attempt. The following is the list of papers which experiments on using CNNs on the CIFAR-10 and CIFAR-100 dataset:

1. Maxout Networks
2. Network in Network
3. Deeply-Supervised Nets
4. Regularization of Neural Networks using DropConnect
5. ImageNet Classification with Deep Convolutional Neural Networks

Pick a paper based on $mod(RollNo., 5) + 1$. You need to discuss the following:

1. What are the novelties introduced in this paper?
2. Illustrate and explain the architecture of the network used?
3. What are the assumptions made regarding the problem to be solved?
4. What are the possible drawbacks? Can you suggest suitable modifications?