Society for Artificial Intelligence and Deep Learning

https://saidl.in

Summer 2019 Induction Assignment

July, 2019

1 About

This is the official Summer assignment for SAiDL inductions. Please fill this google form for registration [Link] and Join our Slack channel [Link] and at the end of the assignment please fill the google form for submission [Link] . If you have any doubts first try to Google it, if still not clear you can pm any one of us. The deadline for the assignment is **8th August 2019**.

2 Getting Started

2.1 Preparing your Machine

- Dual boot your machine with Ubuntu **16.04** LTS
- Install git
- Install the latest version of Anaconda with Python 3.6 (make sure that you agree to change PATH variable when prompted)
- Install OpenCV via python for pip
- Install latest version of Pytorch with or without GPU support depending on your system specs.

2.2 Setting up your Github account

Set up a github account and make sure that you redeem the Student Developer plan [Link].

2.3 General Notes

Before you try and attempt the assignment, it is important that you remember we value your attempt more than the final results. You may or may not have completed the assignment as a whole, but make sure to submit it.

3 Study Resources

3.1 Recommended:

3.1.1 Python 3.6

Python is the Franca Lingua of Ai. You will be learning Python 3.6 . Do it either from “Learn Python 3 the Hard way by Zed Shaw” or tutorial series by sentdex [Link].

3.1.2 Andrew Ng’s Coursera Course on Machine Learning

This course is the launchpad for most Ai enthusiasts and gives you your first hands on approach to Machine Learning along with the maths behind it. You will be implementing ML algorithms in Octave/MATLAB.

3.1.3 Linux Terminal

This is one of the fundamental requirements for any computer scientist. This [Link] will help you get started.

3.1.4 Numpy

Crudely speaking this is MATLAB for Python. I suggest you do this after the Andrew NG course. You can go to its official tutorial or learn it with hands on deep learning experience via deeplearning.ai’s first course.

3.1.5 Pandas

This is one of the most crucial and powerful libraries in data science. To begin with you end to learn how to read and write CSV and JSON files as well how to manipulate Dataframe rows columns and contents. Again sentdex [Link] to the rescue.

3.1.6 Matplotlib

Learn how to plot basic graphs. The pyplot submodule should be enough for the beginning. Sentdex [Link] is your saviour again.

3.1.7 OpenCV

This is one of the best computer vision library out there, However currently we will be using it only to load images as Numpy arrays. Tutorial for the same is here.

3.1.9 Stanford’s CS231n

This is the main course of the entire assignment. You have to see all the lecture videos of Stanford’s Class for Computer Vision using Deep Learning. Make sure you do the Spring 2017 iteration. If possible do the course assignments too. [Youtube link] [Course Link]

3.1.8 Gym

OpenAI's gym is a toolkit for development of reinforcement learning algorithms. It has a lot of agents and environments for which state, action, observation, returns based on actions, etc can generated. You can follow the official documentation.

3.1.9 Mujoco-py

Mujoco-py is a library which together with Gym provides real world physics simulations. Obtain a free student license from [Link]. Apply for this at the earliest as it can sometimes take time for the license key to arrive. After your license key arrives, carefully follow the steps in the install mujoco-py section here [Link].

3.1.10 UCB's CS294

This course is on Deep Reinforcement Learning. Watching just the first two lectures of this series will suffice for doing the imitation learning question in the assignment. [Course Website]

3.1.11 Pytorch Tutorials

You will be doing the assignments in Pytorch and Numpy only. We would suggest that you get hands on experience with Pytorch by following the official tutorials or this series [Link].

3.1.12 LaTeX Tutorials

You will be writing your Paper Review in LaTeX. From writing resumes/CVs to writing research papers and even writing equations in Jupyter Notebook, learning LaTeX helps. Follow the tutorial at [Link] to learn LaTeX. Only the first three modules will be required for this assignment. Use overleaf to create your LaTeX documents.

4 Assignments

There are three questions in the assignment, all three are compulsory. The second question (4.2) has four options (A, B, C, D) from which you have to solve only **ONE** question. The other two questions (4.1, 4.3) have to be done.

4.1 Numpy neural network

In this question you will have to create a neural net in **numpy** which either computes the xor or the xnor of two two bit binary numbers depending on the value of a third input. If the third input is 0, the neural net should compute the xor of the two numbers and if the third input is 1, the neural net should compute the xnor of the two numbers.

4.2 Applied Deep Learning

A. Computer Vision

Your task is Video Classification. The classification will be based on you developing deep architectures using CNNs, RNNs, etc. The dataset you will be using for this is the Youtube Action Classification Dataset (http://www.cs.ucf.edu/~liujg/YouTube_Action_dataset.html). It is a 1GB dataset so can be loaded on google colaboratory or downloaded locally. Many different analysis by researchers have been made around similar datasets. You can take ideas and work on it. For resources you can try to look at attempts at the YouTube 8M dataset on kaggle.

B. Natural Language Processing

Natural Language Processing and Information Retrieval has seen considerable rapid breakthroughs recently.
You need to complete the following two tasks in this question:

- Build a Text-summarization system for summarizing articles on medium. You are free to choose any approach for the text-summarization task. The dataset also has some other data like the article title, number of views/claps, etc. You will have to design the best possible summarizer using all the available data. You can also study your summarized results or compare them with article titles to make comments on your summarizer model's performance and further investigate it. **You'll be judged on your approach and analysis of your model and not only the final result**. Dataset Link: https://www.kaggle.com/hsankesara/medium-articles

- Get familiar with the concepts of **Attention** and **Transformers** in Deep Learning. You can find the lectures for these concepts in the courses CS231n and CS224n. You'll also find some very good blogs for the same. Once you are familiar with these concepts, you'll have to study **any one** of the following SOTA models in detail : **BERT/ XLNet/ GPT-2.** These are fairly complex models and you'll come across various standard benchmarking tasks, datasets and metrics. So you'll learn about these too along with the architecture of these models. You are expected to make a write-up of about 250 words in Latex on the model of your choice after you are done to finish this task.

C. Speech

Speech processing and generation is a very important field of Deep Learning. Speech data is much more intricate than text data due to the additional information it carries like the tone, pitch, varying amplitude, etc other than the content. Hence it is much more complicated to analyze.

In this question you will be classifying speech based on the content in the audio. Dataset link : https://github.com/Jakobovski/free-spoken-digit-dataset. This dataset is also popularly known as Speech MNIST as it contains voices of people speaking numbers from 1 to 10.

You are required to make a Deep Neural Network architecture using Pytorch and train it to do digit classification using the above **speech** data. After the classification, you are required to do clustering using the last layer of your neural network(No hard and fast rule to only use the last layer for clustering). You can use t-SNE or other dimensionality reduction algorithms to visualize your clustering results.

Again, you will be assessed on your attempt, rather than just on the accuracy of your classification or how good your clustering looks. You can submit it as a Jupyter Notebook or a python script but ensure to document everything you do in detail. Bonus points for anyone who does some interesting exploratory analysis on the dataset/final clustering results.
Resources:

1) https://github.com/mdda/cnn-speech-mnist
2) https://github.com/adhishthite/sound-mnist

D. Reinforcement Learning

This question will have two parts:

- You have to model a Deep Q-network for learning Ms Pacman. One of the key areas where Reinforcement Learning shines is in its applications to games. In order to apply RL to any task however we need to figure out how to formulate the task appropriately(which you will be doing in Part 2). Fortunately gaming already has existing environments in Open AI Gym waiting to be leveraged. Link : https://gym.openai.com/envs/MsPacman-v0/

  Your goal is to design a model based on Q-learning approaches. You can propose modifications to the basic components like the loss function used or the way action exploration is done.

  To make things easier we have also provided starter code which already includes a valid preprocessing module for the agent and a cell to help you familiarize yourself with the environment. Clearly document any changes you make to the starter code. You are also encouraged to play around with the hyperparameters, although heavy deviations from the suggested ones in the starter code are not recommended.

  Starter Code link:
  https://colab.research.google.com/drive/19_E9yyJtJ3jXBxNLi2G_yH5XI2ssJISf

- You receive the following letter:

  Dear Friend,
  Some time ago, I bought this old house, but found it to be haunted by ghostly sardonic laughter. As a result it is hardly habitable. There is hope, however, for by actual testing I have found that this haunting is subject to certain laws, obscure but infallible, and that the laughter can be affected by my playing the organ or burning incense. In each minute, the laughter occurs or not, it shows no degree. What it will do during the ensuing minute depends, in the following exact way, on what has been happening during the preceding minute: Whenever there is laughter, it will continue in
  the succeeding minute unless I play the organ, in which case it will stop. But continuing to play the organ does not keep the house quiet. I notice, however, that whenever I burn incense when the house is quiet and do not play the organ it remains quiet for the next minute. At this minute of writing, the laughter is going on. Please tell me what manipulations of incense and organ I should make to get that house quiet, and to keep it so.
  Sincerely,
  At Wits End

a) Formulate this problem as an MDP (for the sake of uniformity , formulate it as a continuing discounted problem with γ = 0.9. Let the reward be +1 on any transition into the silent state, and -1 on any transition into the laughing state). Explicitly give the state set, action sets, state transition and reward function.

b) Starting with the policy π(laughing) = π(silent) = π(incense, no organ), perform a couple of policy iterations(by hand) until you find an optimal policy(Clearly show and label each step. If you are taking a lot of iterations, stop and reconsider your formulation). Do a couple of value iterations as well.

c) What are the resulting optimal state-action values for all state-action pairs?

d) What is your advice to At Wits End?

Write your answers in a LaTeX file, create a pdf and upload it to your Github repository containing the rest of your code.

4.3 Paper Review

In this question your task is to review a deep learning paper from the papers listed here. First read the abstract and give a rough reading of the paper you have chosen. Identify the problem that the paper aims to solve. Also identify important details in the paper including important assumptions, models, results, etc. Lastly describe these details and write a brief review in about 300 words. The review should be in LaTeX. Use overleaf for the same.

The papers are:

1) https://arxiv.org/abs/1506.02216
2) https://nlp.stanford.edu/pubs/sidaw13fast.pdf
3) https://arxiv.org/abs/1703.03400
4) https://arxiv.org/abs/1803.10122

6. Submission

After completing the assignment (3 questions) upload your code on Github in a repository titled SAiDL-Summer-Assignment-2019. The Deadline is 8th August 2019. You can learn how to use git from codeacademy [Link]. To submit fill this google form [Link]. Inform us when you complete each question from the assignment (submit the google form only after completion of the whole assignment or after deadline).

7. Contact Info

General doubts or doubts that you think might help others as well should be asked on #general channel. Doubts regarding the first and third question can be asked to anyone. If you have any doubts pertaining to a specific question, then reach out to the person(s) assigned for that task (check details below) first via pm.

Our Team -
Sharad Chitlangia
Rijul Ganguly
Pranav Mahajan (RL)
Ajay Subramanian (RL)
Souradeep Chakraborty (RL)
Rajaswa Patil (NLP)
Het Shah (CV)
Ashwin Vaswani (CV)
Mehul Rastogi (Speech)