**Homework 1**

Due: 03/01/2021 11:59PM EST

**Introduction**

This homework will involve you developing several ML methods to classify the most popular benchmark computer vision datasets. You will be given two real-world dataset: MNIST, CIFAR-10. All detailed descriptions are online. You need to build traditional ML models and deep neural networks. These classifiers will be further employed in future assignments to study adversarial attacks and defenses.

**Note that,** you can use all online resource to finish your homework, but you need to learn and have your own understandings, and I will random pick up and ask your understandings from our homework evaluation.

**Setup**:

**Environment**: You can use any coding environments.

**Code Hints and examples**: same as this tutorial: <https://pytorch.org/tutorials/beginner/blitz/cifar10_tutorial.html>

<https://www.tensorflow.org/tutorials/images/cnn>

**Dataset**:

Since they are so popular online, you can download them directly from online, e.g., <http://yann.lecun.com/exdb/mnist/>

<https://www.cs.toronto.edu/~kriz/cifar.html>

**Tasks – Computer Vision Image Classification:**

You are going to implement the most popular frameworks on the benchmark computer vision datasets. The different tasks bring you more understanding between learning frameworks and datasets.

*For each task, you will do (1) data load (2) model setup (3) model training (4) model evaluation (5) reload model without training for evaluation. I will check all of them and ask questions during our homework evaluation course section.*

**Task #1 [Train Traditional ML] [5 points]:**

Decision Tree and other traditional machine learning to these datasets.

**Hint**: you can use sk-learn package, choose **3-5** methods to above datasets: https://scikit-learn.org/stable/modules/generated/sklearn.ensemble.GradientBoostingClassifier.html

**Task #2 [Train NN] [5 points]:**

Two layers NNS: one hidden layer and one output layer.

**Task #3 [Train CNN] [5 points]:**

2-5 layers CNNS: see the tutorials. **Hints**, for MNIST, you can use two layers, CIFAR, you need more layers.

**Bonus Task [Gradient extraction] [1 points]:**

Given a data sample and extract the gradient information from the data. If you do not do it now, you still need to use this technology to finish the second homework.

**Grading Policy**

1. A clear report what you done and all results: *(1) data load (2) model setup (3) model training (4) model evaluation (5) reload model without training for evaluation*
2. Clearly describe your architecture and what you have tried.
3. You do not need to train and find the best model. But you need to report your thinking during your training tasks.
4. All grading would be done during evaluation class. Prepare your trained model, make sure you can train and reload it and show the same results as you reported.

I wish all students can better understanding deep learning through various datasets and tasks. **No delay for this homework in general.**

A simple table report for traditional ML:

Decision Tree Naive Bayesian SVM etc

MNIST xxx xxx xxx xxx

CIFAR xxx xxx xxx xxx

A simple table report for DNNs:

Two layer Three layer Four layer Five layer ResNet-x (optional)

MNIST xxx xxx xxx xxx xxx

CIFAR xxx xxx xxx xxx xxx

**A better model based on your better understanding!**