This homework is designed to test your pre-requisite knowledge of Machine Learning and Deep Learning. Please submit your homework using this google form link: <https://forms.gle/t2bjNorQYbbjAhBH6>

Deadline for submission**: September 18, 2020**

Please do not share this homework outside of class.

Points**:** Questions 1-7 are 1 point each. Question 8 is 5 points. Total: 12 points.

1. Consider the following training set:

|  |  |
| --- | --- |
| **x** | **y** |
| 1 | 0.5 |
| 2 | 1 |
| 4 | 2 |
| 0 | 0 |

If we are using a linear regression model. What would the final model be after fitting it on this data?

2. What is K-Means. How is it different from Meanshift?

3. What is softmax? Under what cases does softmax computation become unstable?

4. What is regularization in CNNs? Explain 1 common technique used for regularization.

5. What is a GAN? What does it mean for a GAN to ‘mode collapse’?

6. Describe how you would compute PCA for a set of feature data points?

7. What are some common initialization methods for Convolution layers and Fully Connected layers in a CNN?

8. One interesting application of machine learning is in the area of medical diagnoses. Implement Neural Network in python to classify the data into Benign or Malignant for the Wisconsin Diagnostic Breast Cancer (WDBC) dataset.

Attach the code and the output.

Things to keep in mind:

How will you choose the features?

How to overcome overfitting?

How many layers should be there?

Dataset: <http://archive.ics.uci.edu/ml/datasets/Breast+Cancer+Wisconsin+(Diagnostic)>

Download Data from here: http://archive.ics.uci.edu/ml/machine-learning-databases/breast-cancer-wisconsin/breast-cancer-wisconsin.data

Dataset Description:

# Attribute Domain

-------------------------------------------------------------------------------

1. Sample code number id number

2. Clump Thickness 1 - 10

3. Uniformity of Cell Size 1 - 10

4. Uniformity of Cell Shape 1 - 10

5. Marginal Adhesion 1 - 10

6. Single Epithelial Cell Size 1 - 10

7. Bare Nuclei 1 - 10

8. Bland Chromatin 1 - 10

9. Normal Nucleoli 1 - 10

10. Mitoses 1 - 10

11. Class: (2 for benign, 4 for malignant)

Please attach training curves, final losses and visualizations of errors made by the trained model, along with a description of the methods/tricks you tried to improve performance.