COL774: Machine Learning **Date: April 7, 2019**

Assignment 3 Report

Name: Manish Tanwar **Entry Number: 2016CS10363**

1. Decision Trees:

Part (a):

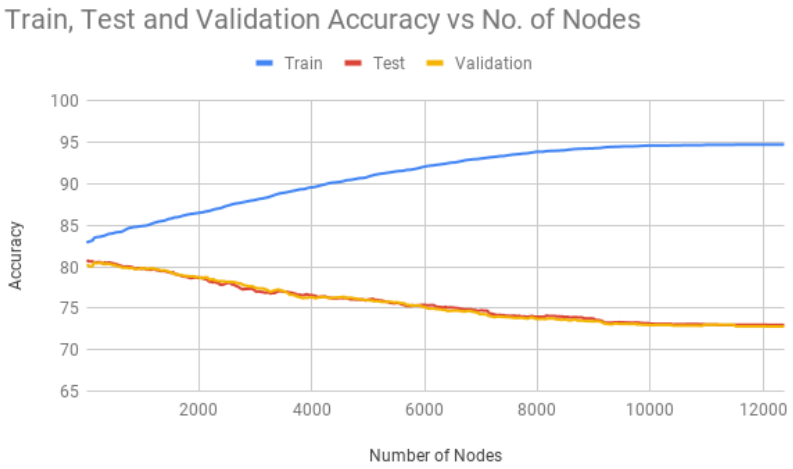


Figure 1: Accuracy Vs No. of Nodes

Part (c):

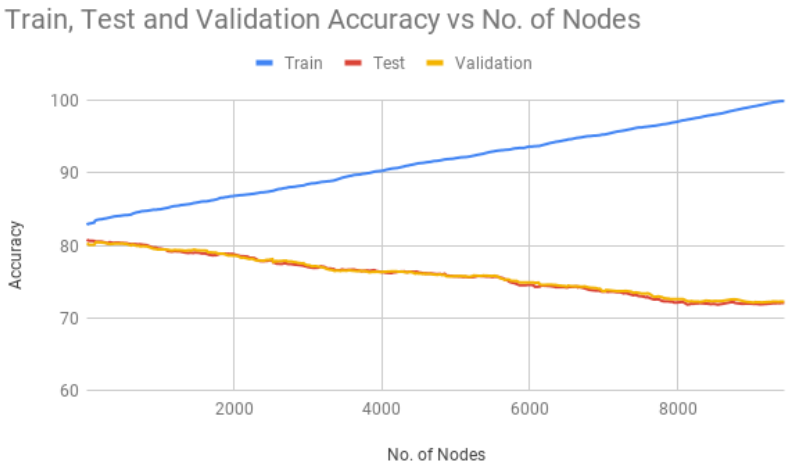


Figure 2: Accuracy Vs No. of Nodes

Multisplit Atrributes: Following attributes split multiple times along a branch with their maximum number of split along a path from root to node.

Atrribute	Max No. of Split
X1	6
X5	4
X12	4
X13	4
X14	3
X15	3
X16	3
X17	3
X18	4
X19	4
X20	3
X21	3
X22	3
X23	3

Part (d):

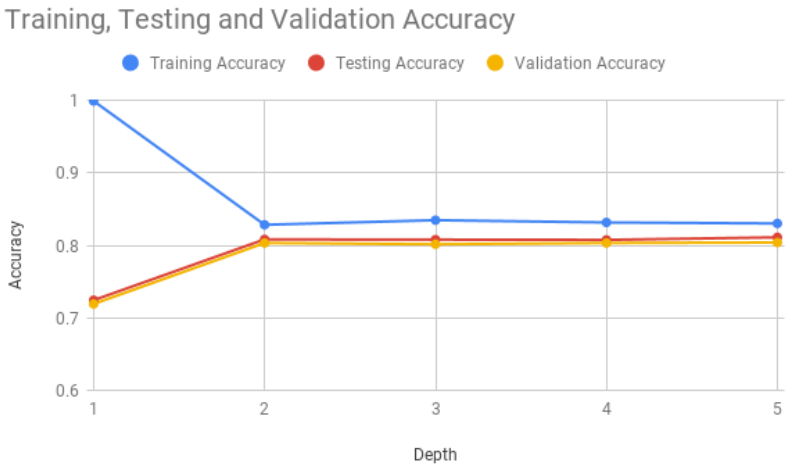


Figure 3: Accuracy Vs Depth

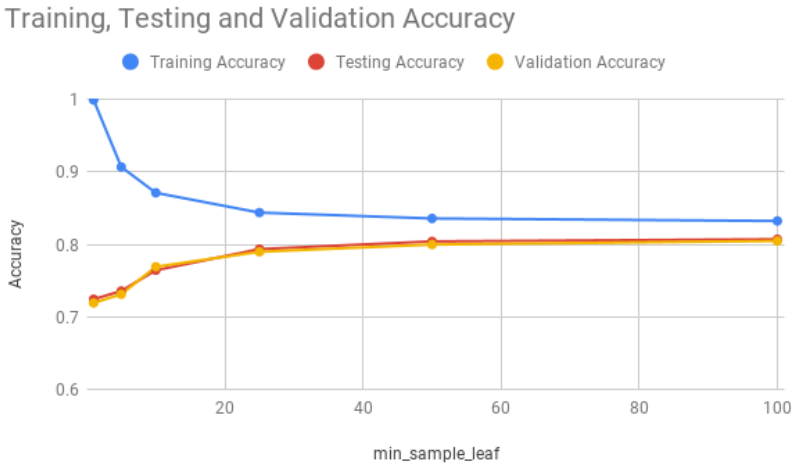


Figure 4: Accuracy Vs Min sample leaf

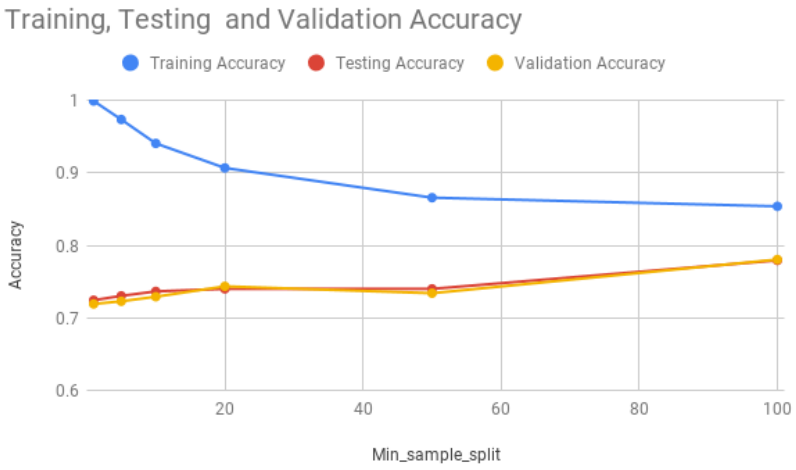


Figure 5: Accuracy Vs Min Sample Split

- **Best Parameters:** max_depth = 4, min_samples_split = 10, min_samples_leaf = 30
- **Accuracies Obtained:** Test Accuracy: 80.85%, Train Accuracy: 83.20%, Validation Accuracy: 80.52 %

Part (e):

- **Best Parameters:** max_depth = 50, min_samples_split = 100, min_samples_leaf = 50
- **Accuracies Obtained:** Test Accuracy: 80.71%, Train Accuracy: 83.18%, Validation Accuracy: 80.45 %

Part (f):

- **Best Parameters:** n_estimators=120, max_features=2, bootstrap = True
- **Accuracies Obtained:** Test Accuracy: 79.56%, Train Accuracy: 86.63%, Validation Accuracy: 79.51 %

2. Neural Network:

Implementation Specifications:

- Used Softmax activation function in the output layer.
- Used cross entropy loss as loss funciton.

Part (a):

One-hot Encoding Dataset Link: https://drive.google.com/drive/folders/1VVqfcNAuw_i5A9KtQhZJur_dFKMND5h1?usp=sharing

Part (c):

- **Stopping Criteria:**
  Stopping criteria is either of these two condition is true:
  1. $\left|Loss(\theta)^{(t+1)} - Loss(\theta)^{(t)}\right| < \epsilon$ (for a sufficiently small ϵ (took $\epsilon = 10^{-7}$))
     where $Loss(\theta)^t$ is the loss after t^{th} epoch.
  2. Constrainted on number of epochs with 1000.

- **Accuracy and Trainig Time:**

Hiddern Layer Units	Training Time	Training Accuracy	Testing Accuracy
5	107 sec	57.24%	52.38%
10	122 sec	58.7%	58.5%
15	147 sec	58.8%	54%
20	175 sec	96.22%	93.93%
25	185 sec	97.72%	95.17%

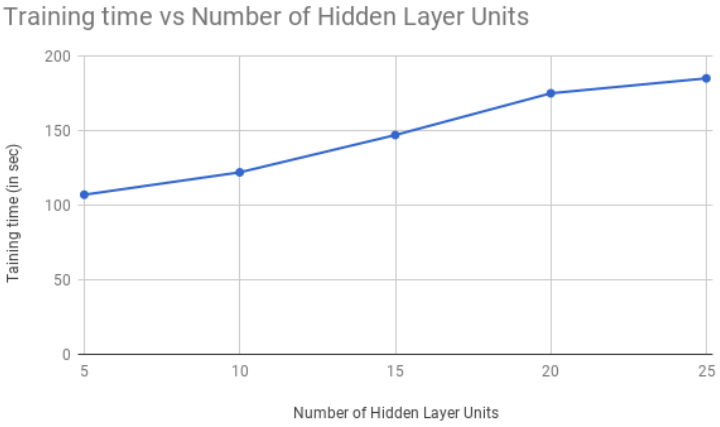


Figure 6: Training time vs No. of Hidden Layer Units

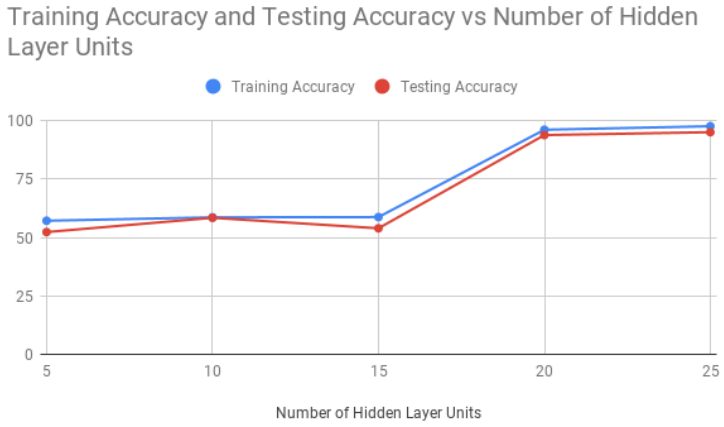


Figure 7: Training & Testing Accuracy vs No. of Hidden Layer Units

- **Confusion Matrices:**

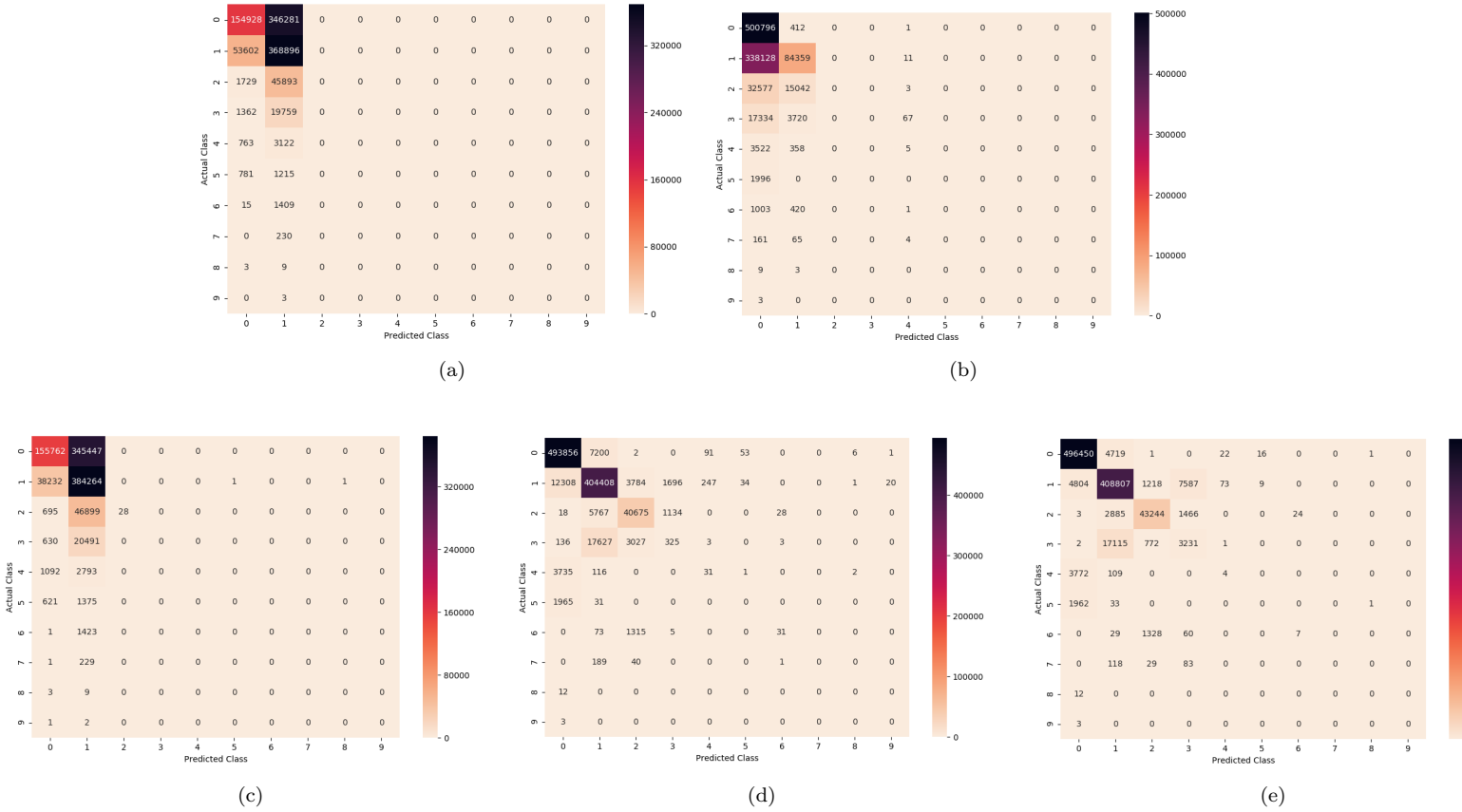


Figure 8: Confusion Matrices for hidden layers units = [5,10,15,20,25]

Part (d): Double Layer:

- **Accuracy and Trainig Time:**

Hiddern Layer Units(for both layers)	**Training Time**	**Training Accuracy**	**Testing Accuracy**
5	158 sec	57.14%	55.70%
10	194 sec	79.69%	78.56%
15	239 sec	90.62%	89.38%
20	294 sec	98.37%	96.87%
25	397 sec	99.26%	96.30%

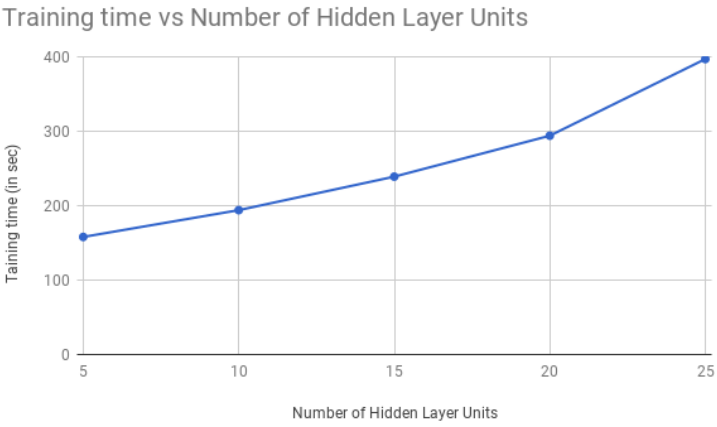


Figure 9: Training time vs No. of Hidden Layer Units

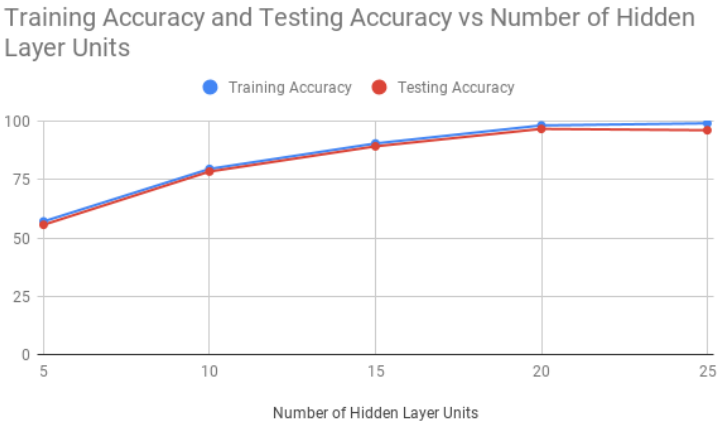
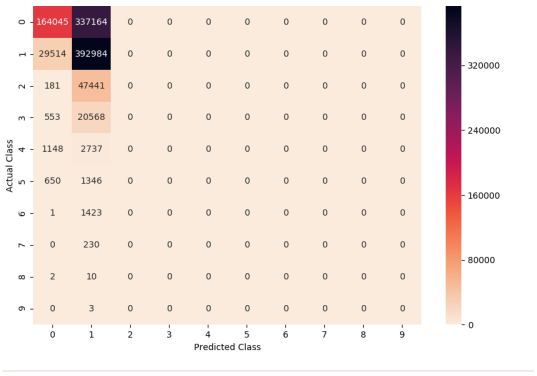


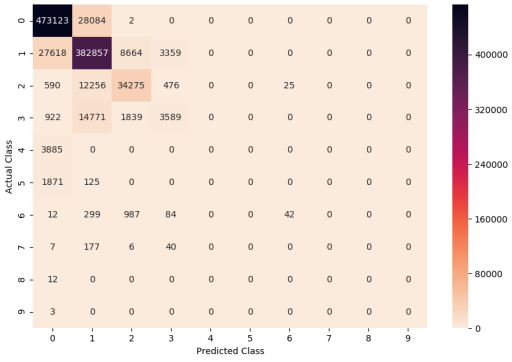
Figure 10: Training & Testing Accuracy vs No. of Hidden Layer Units

- **Confusion Matrices:**

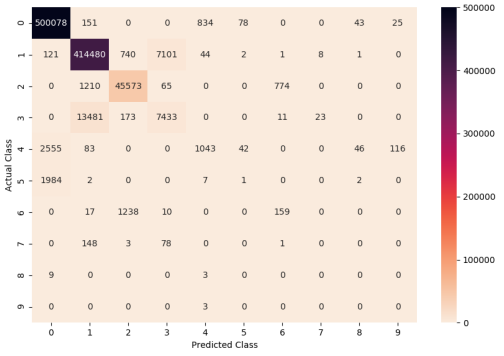


(a)

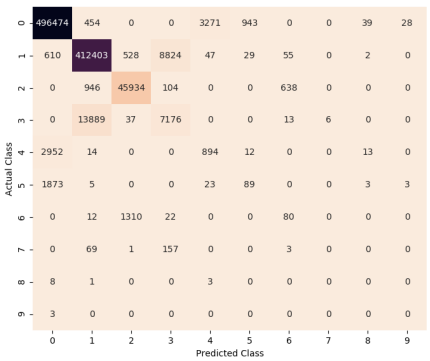
(b)



(c)



(d)



(e)

Figure 11: Confusion Matrices for hidden layers units = [5,10,15,20,25] for both layers

Part (e):

(i) Single Layer:

- **Accuracy and Trainig Time:**

Hiddern Layer Units	Training Time	Training Accuracy	Testing Accuracy
5	6.09 sec	54.10%	54.17%
10	8.7 sec	73.02%	71.56%
15	24 sec	79.47%	77.62%
20	31 sec	95.44%	94.5%
25	20.14 sec	95.04%	93.53%

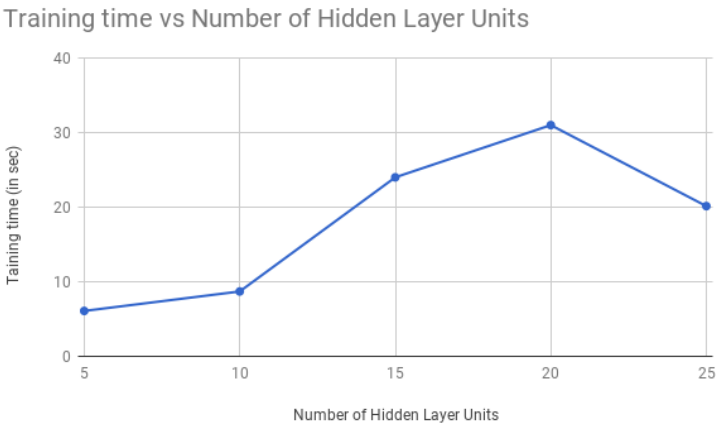


Figure 12: Training time vs No. of Hidden Layer Units

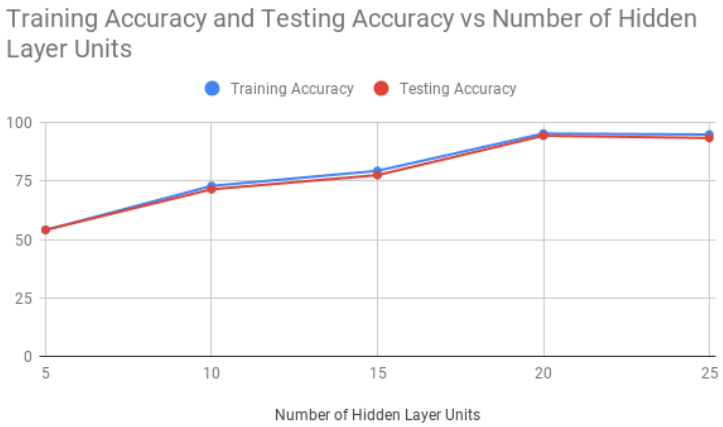
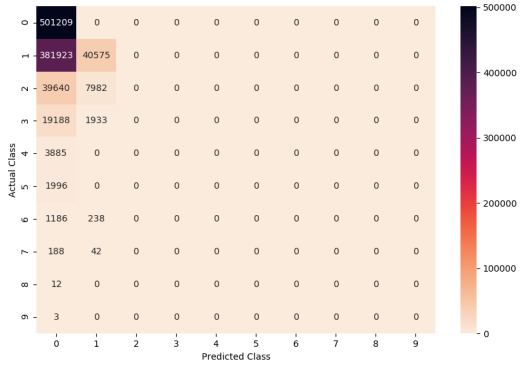
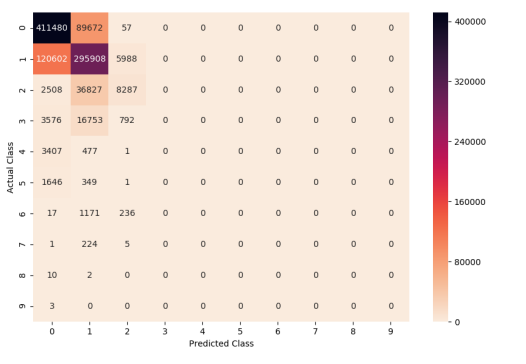


Figure 13: Training & Testing Accuracy vs No. of Hidden Layer Units

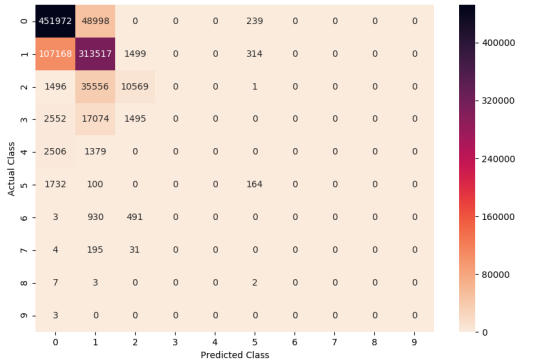
- **Confusion Matrices:**



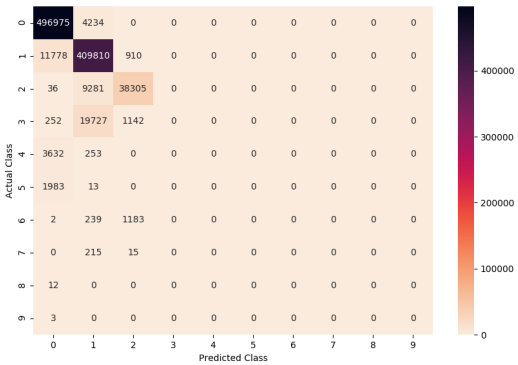
(a)



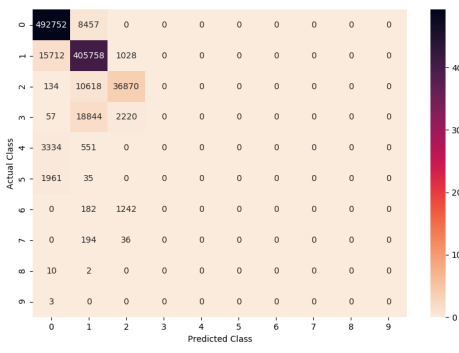
(b)



(c)



(d)



(e)

Figure 14: Confusion Matrices for hidden layers units = [5,10,15,20,25]