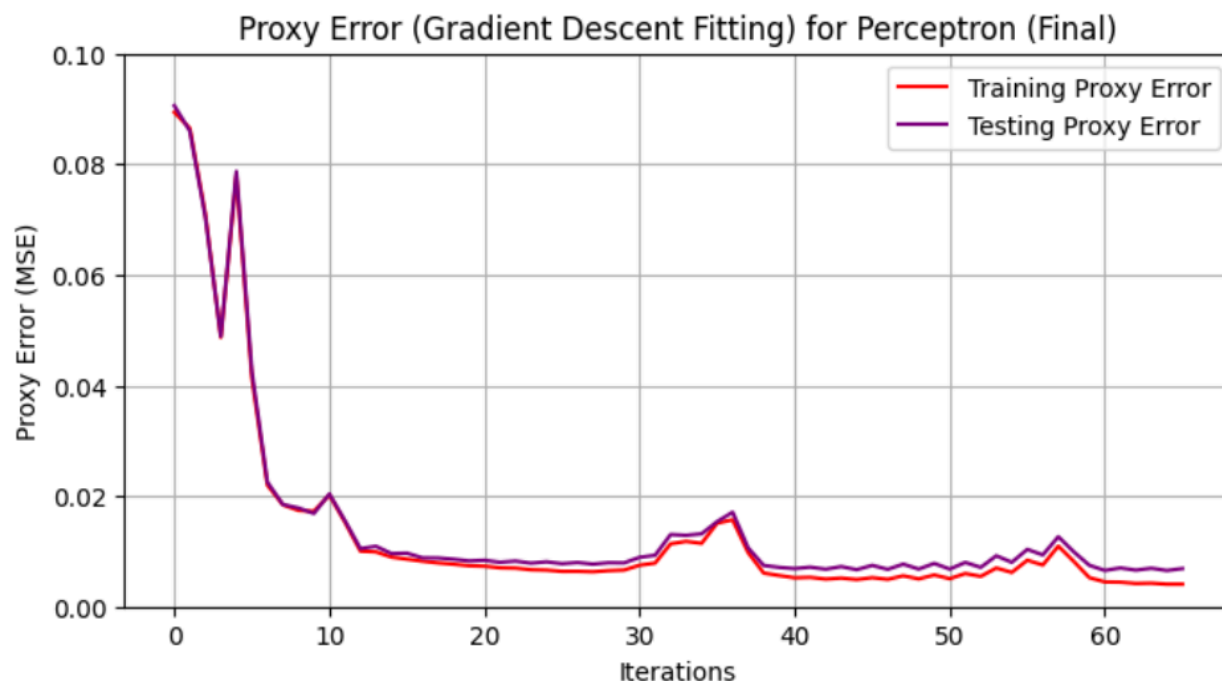


## Perceptron Model:

### *Gradient Descent Fitting Curves:*

#### *a) Training and Testing Proxy Errors:*

This graph shows how the training and test proxy errors (mean squared error) changed across iterations of gradient descent for the final Perceptron model. I used a learning rate of 16, which was the best value found during cross-validation. There's a steep drop in error early on, which tells me the model was learning quickly. After that, both the training and testing errors level off with just small oscillations, which suggests that the model reached convergence and is generalizing well to unseen data. The numbers shown in the screenshot below are just the output values being plotted.



## *CIS 481 Project 2: Handwritten Digits Recognition using Neural Networks*

*Written by: Zaynab Mourtada*

```
>>> Perceptron (Final) - ALL Training Proxy Error:
[0.08940029 0.08649373 0.07100457 0.048709 0.07793545 0.04141182
0.02199365 0.0184237 0.0174094 0.01725697 0.02019469 0.01528627
0.01003728 0.00990745 0.00896484 0.00859677 0.00822444 0.00791586
0.00768912 0.00740567 0.00732436 0.00702193 0.00697364 0.00667646
0.00661142 0.00637903 0.00638701 0.00631024 0.00650759 0.00662119
0.00749795 0.00788094 0.01136874 0.01181166 0.01145384 0.01510496
0.01570356 0.00978496 0.00608322 0.00563119 0.00522276 0.00530579
0.00499566 0.00517476 0.00491425 0.00524955 0.00494896 0.00558249
0.00502991 0.00572816 0.0050645 0.0059908 0.00548141 0.00699028
0.00619883 0.00839892 0.00753487 0.01095223 0.00817758 0.00524853
0.00447405 0.00443205 0.00420349 0.00425547 0.00407476 0.00407493]

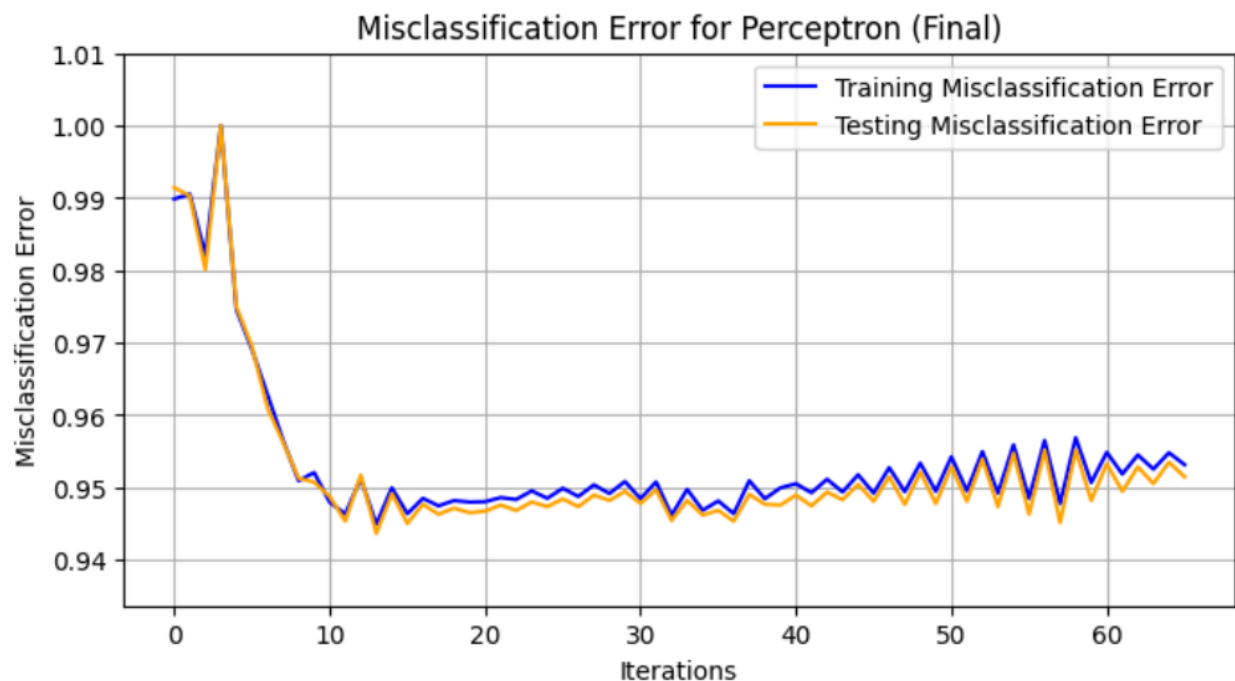
>>> Perceptron (Final) - ALL Test Proxy Error:
[0.09059579 0.08601736 0.07048576 0.04888932 0.07870106 0.04256798
0.02252682 0.01847131 0.01784927 0.01685971 0.02038009 0.01556792
0.01049712 0.01092049 0.00960825 0.00966902 0.00883224 0.00880855
0.00858557 0.00826814 0.00838611 0.00802858 0.00824224 0.0078735
0.00811179 0.00777825 0.00798748 0.00769426 0.00795124 0.00791697
0.00891852 0.00929391 0.01304249 0.01290207 0.01317719 0.01538205
0.01708556 0.01062737 0.00745944 0.00710594 0.00689113 0.00714995
0.00678016 0.00725121 0.00669147 0.00746817 0.0067533 0.00769144
0.00679893 0.00782142 0.00679665 0.00799949 0.00710787 0.00920152
0.0080238 0.01036 0.00935925 0.01265057 0.00990871 0.00748483
0.00654933 0.00700355 0.00660888 0.00695014 0.00654142 0.00689008]
```

### *b) Training and Testing Misclassification Errors:*

This plot tracks how the misclassification error evolved during training for the Perceptron model. Again, I used the best learning rate of 16. There's a sharp drop in error early on, and then it stabilizes around 0.95. What stands out is how closely the training and test curves follow each other; that tells me the model isn't overfitting and maintains consistent performance. As indicated before, the numbers in the screenshot below simply represent the values that were plotted on the graph.

## CIS 481 Project 2: Handwritten Digits Recognition using Neural Networks

Written by: Zaynab Mourtada



```
>>> Perceptron (Final) - ALL Training Misclassification Error:
[0.98986556 0.99053775 0.98210962 1.          0.9745605  0.96907963
 0.96282316 0.95651499 0.95093071 0.95206825 0.94798345 0.94632885
 0.95124095 0.94493278 0.94994829 0.94638056 0.94850052 0.94746639
 0.94819028 0.94798345 0.94803516 0.94860393 0.9483454  0.94953464
 0.94850052 0.94989659 0.94875905 0.95036194 0.9491727  0.9508273
 0.94844881 0.95077559 0.94617373 0.94974147 0.94684592 0.94813857
 0.94638056 0.95093071 0.94844881 0.94994829 0.95051706 0.94927611
 0.95113754 0.94937952 0.95175801 0.94922441 0.95274043 0.94943123
 0.95336091 0.94948294 0.95423992 0.94953464 0.95496381 0.94922441
 0.95589452 0.94850052 0.95651499 0.94777663 0.95687694 0.95067218
 0.95486039 0.95186143 0.95449845 0.95253361 0.95480869 0.95315408]

>>> Perceptron (Final) - ALL Test Misclassification Error:
[0.99143763 0.99027484 0.98012685 1.          0.97494715 0.96945032
 0.96099366 0.95623679 0.9512685  0.95073996 0.9487315  0.94534884
 0.95169133 0.94365751 0.94915433 0.94503171 0.94767442 0.94630021
 0.94714588 0.94651163 0.94672304 0.94756871 0.94682875 0.94799154
 0.94735729 0.94841438 0.94735729 0.94894292 0.94820296 0.94947146
 0.94778013 0.94968288 0.94545455 0.94820296 0.9461945  0.94682875
 0.94534884 0.94904863 0.94767442 0.94756871 0.94894292 0.947463
 0.94936575 0.94830867 0.95042283 0.94809725 0.95147992 0.94767442
 0.95211416 0.94778013 0.95285412 0.94809725 0.95391121 0.94735729
 0.95475687 0.94630021 0.9551797  0.94513742 0.9551797  0.94820296
 0.95327696 0.94947146 0.95285412 0.95052854 0.95348837 0.95147992]
```

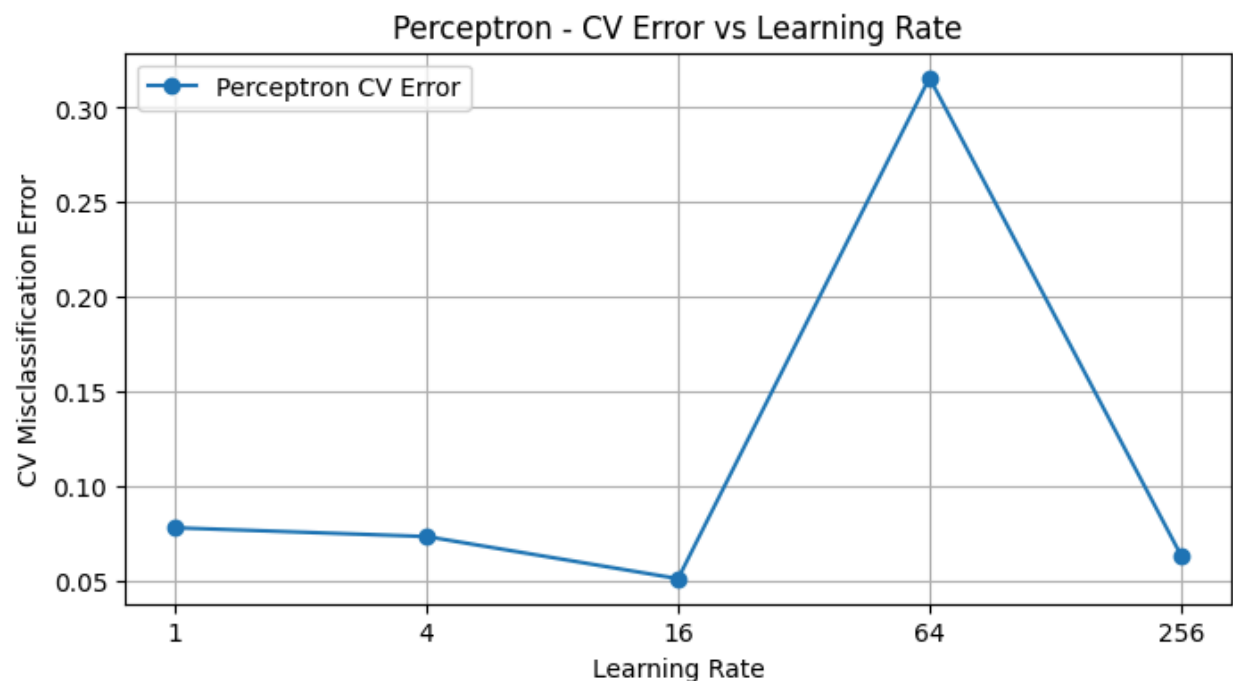
### ***c) Evaluation of Trial Data Set***

The perceptron had a 0% misclassification error rate on the trial set, which is a bit unexpected given how simple the model is. One possible reason is that the trial set is small and may not be complex enough to expose the perceptron's limitations. It's also likely that the trial examples happen to align well with the model's single-layer decision boundary, making it easier to separate the classes without needing deeper feature learning.

### ***CV Misclassification Errors:***

Based on 3-fold cross-validation, the best-performing learning rate for the Perceptron model was determined to be 16.

Learning Rate	Average CV Error
1	0.0781
4	0.0734
16	0.0512
64	0.3158
256	0.0636

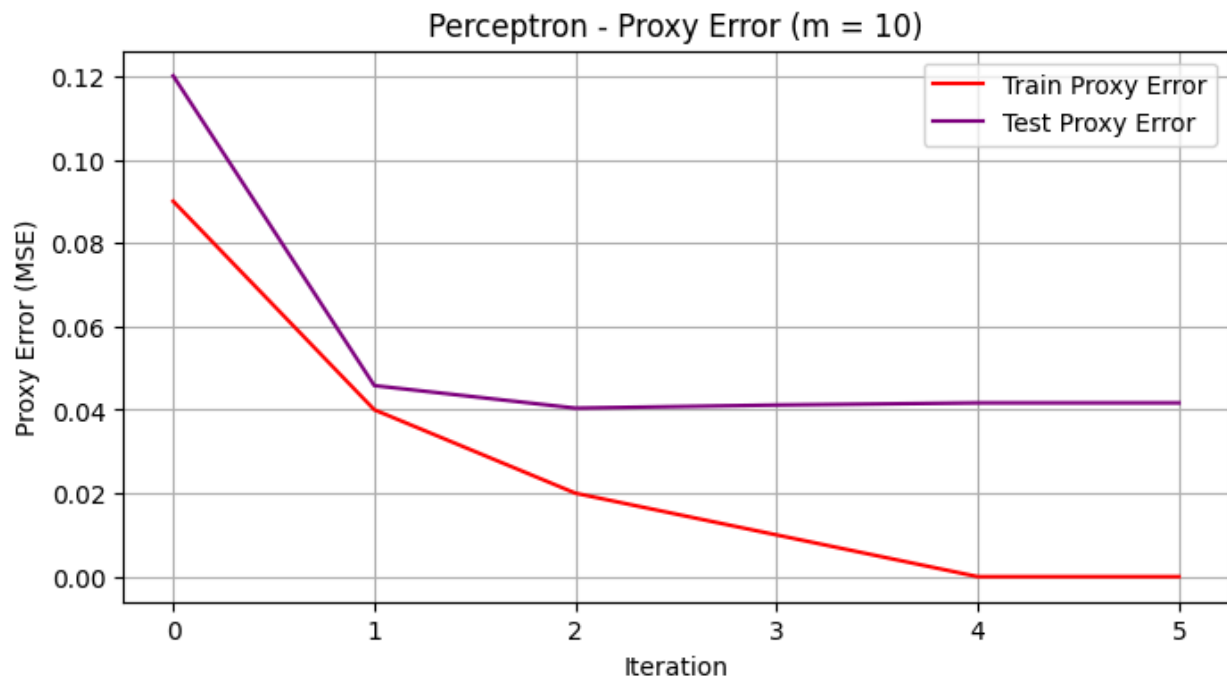


## ***Learning Curves:***

### ***a) Learning Curves by Subset Size (Proxy & Misclassification Error vs Iterations)***

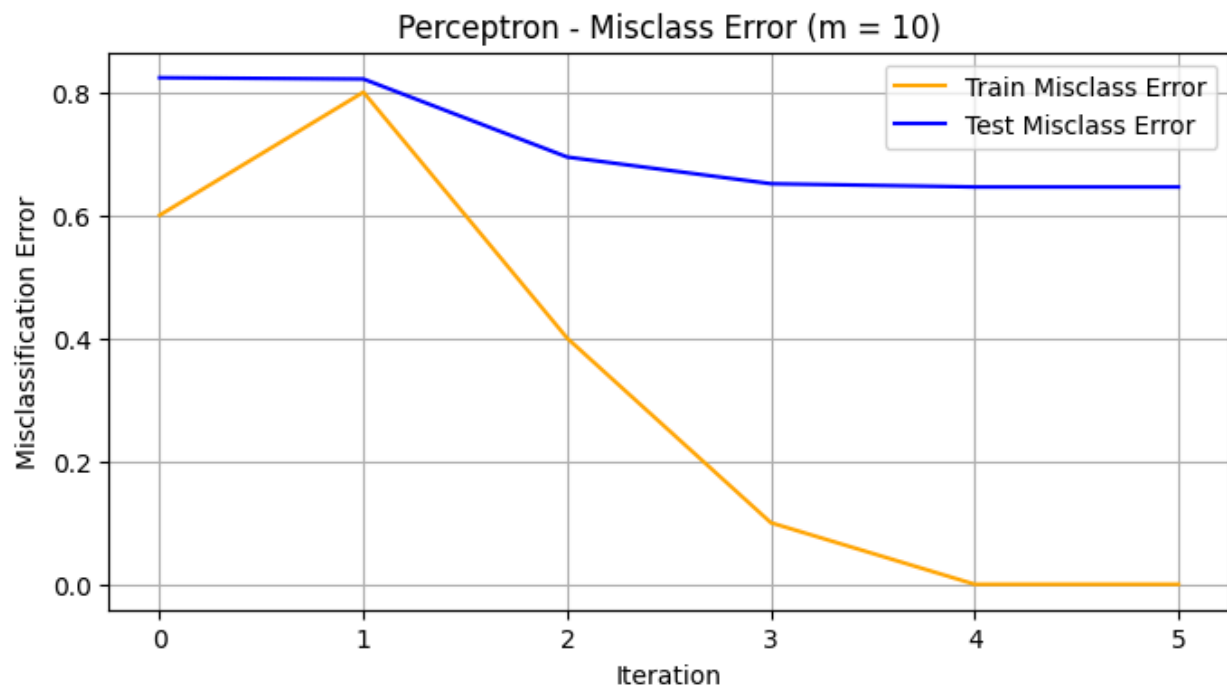
The following graphs show the learning curves (proxy and misclassification error) across iterations for each subset size  $m$ , highlighting both training and testing performance.

- $m = 10$

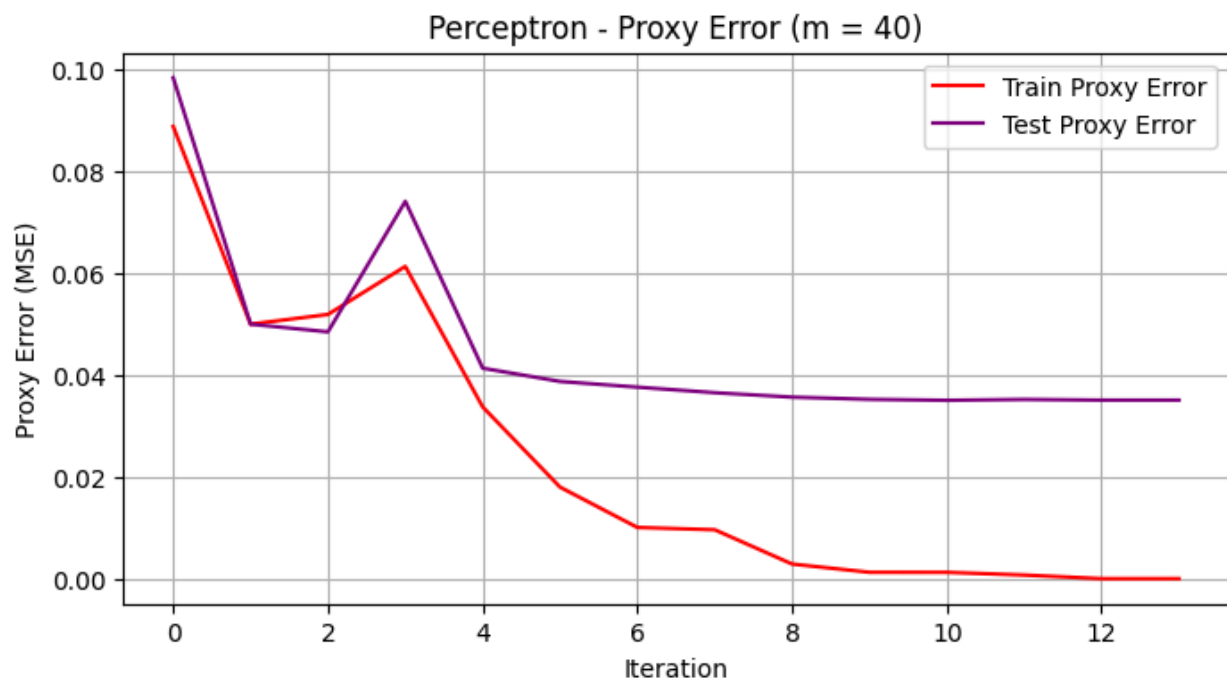


## CIS 481 Project 2: Handwritten Digits Recognition using Neural Networks

Written by: Zaynab Mourtada

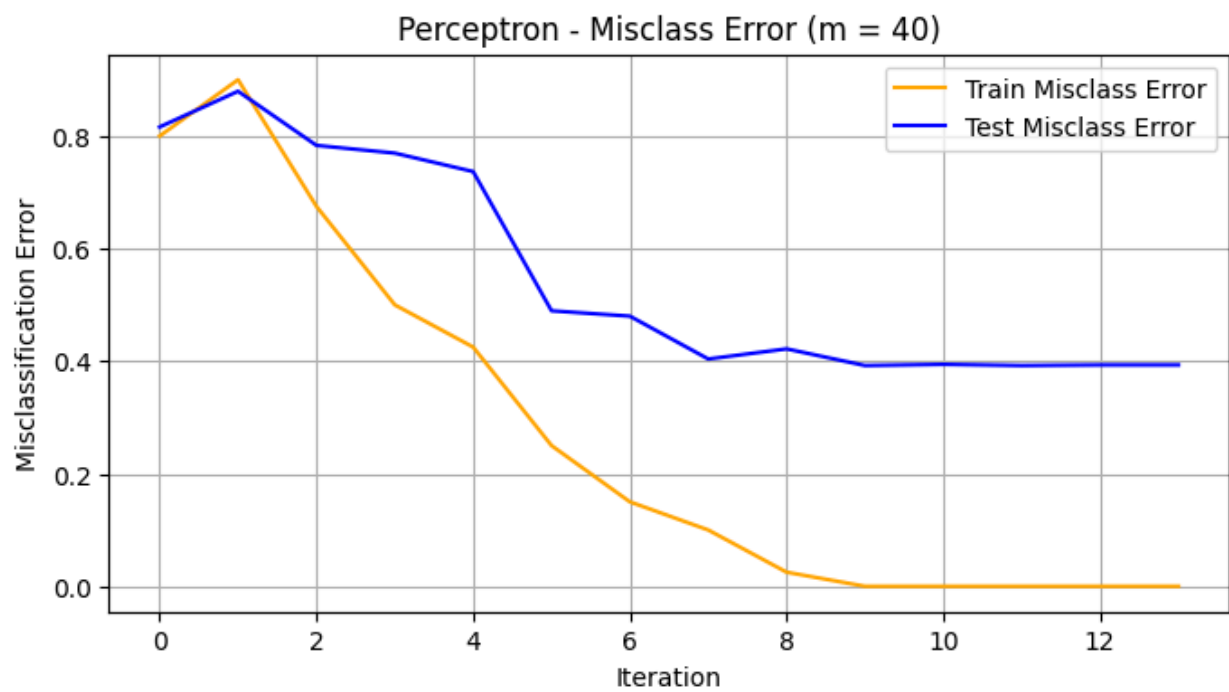


- m = 40

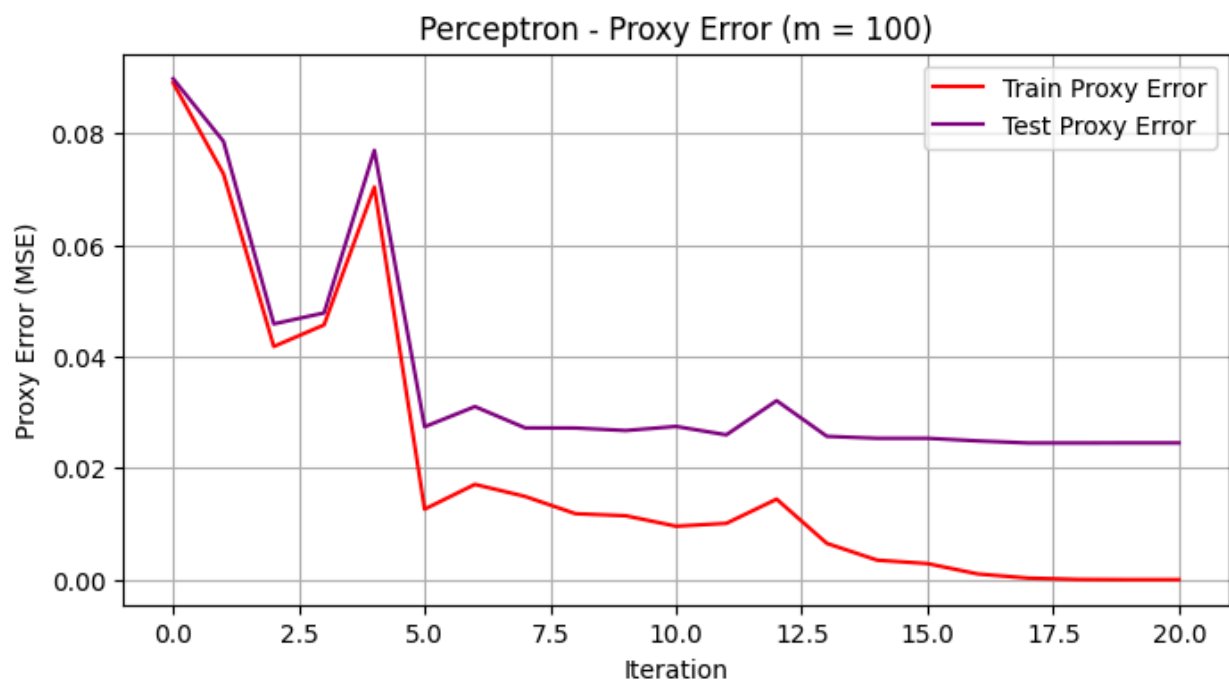


**CIS 481 Project 2: Handwritten Digits Recognition using Neural Networks**

**Written by: Zaynab Mourtada**

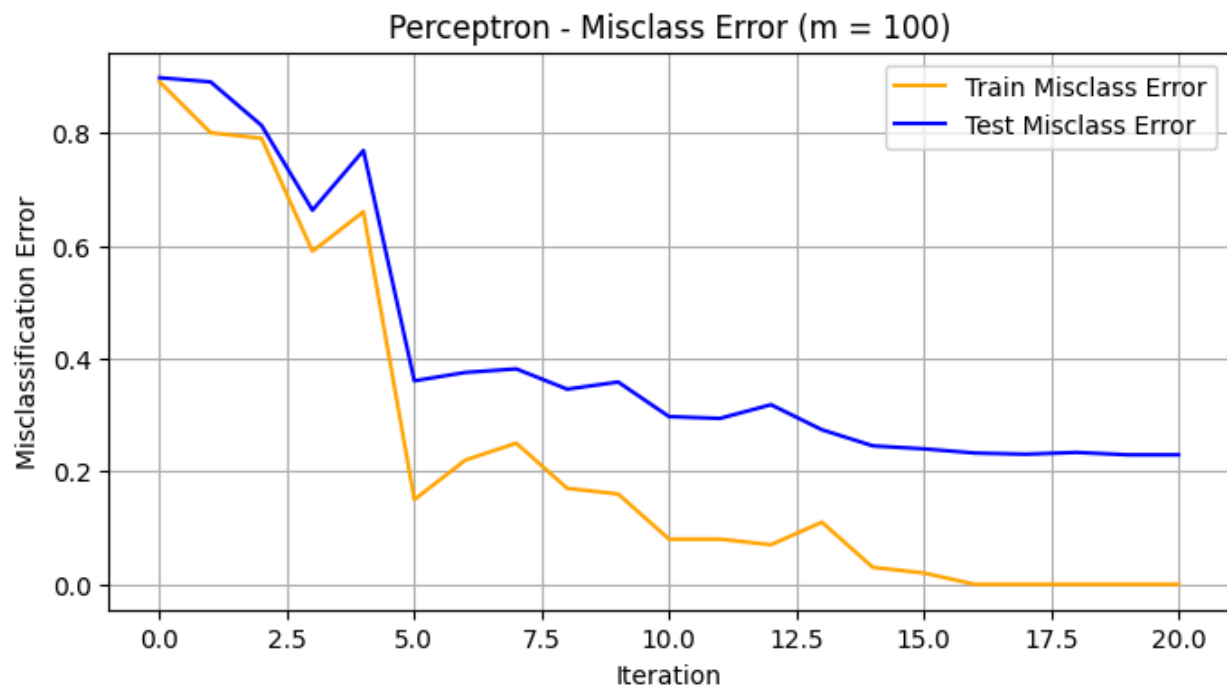


- 
- m = 100

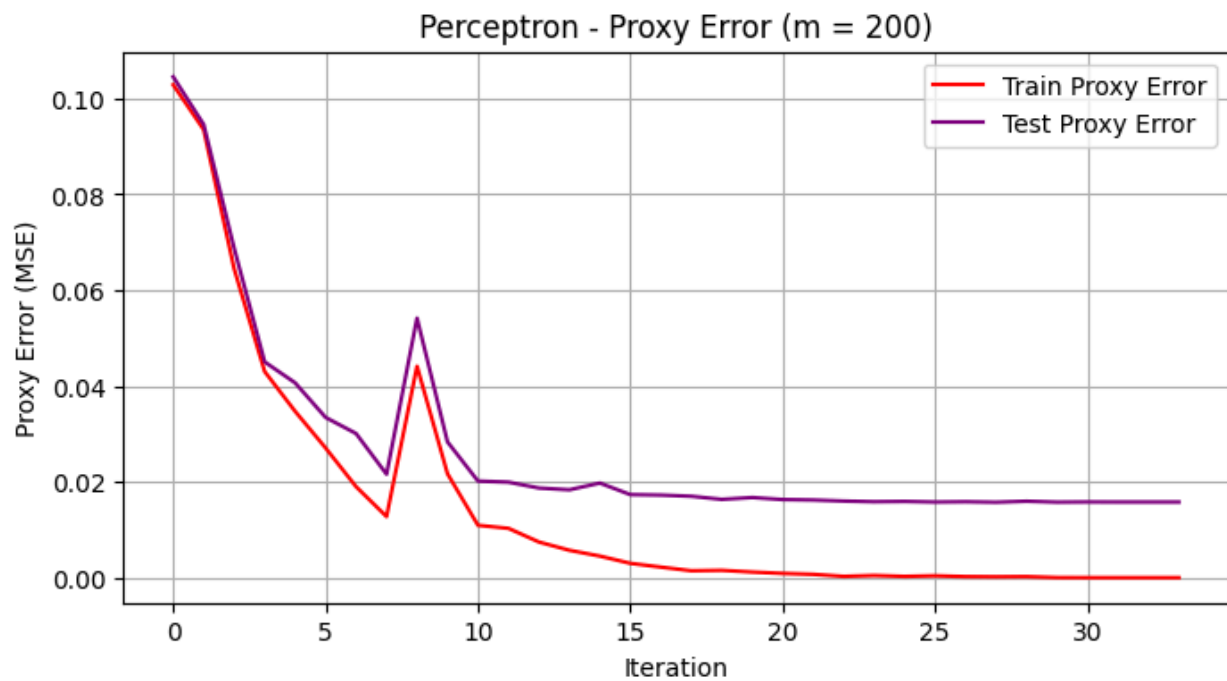


## CIS 481 Project 2: Handwritten Digits Recognition using Neural Networks

Written by: Zaynab Mourtada



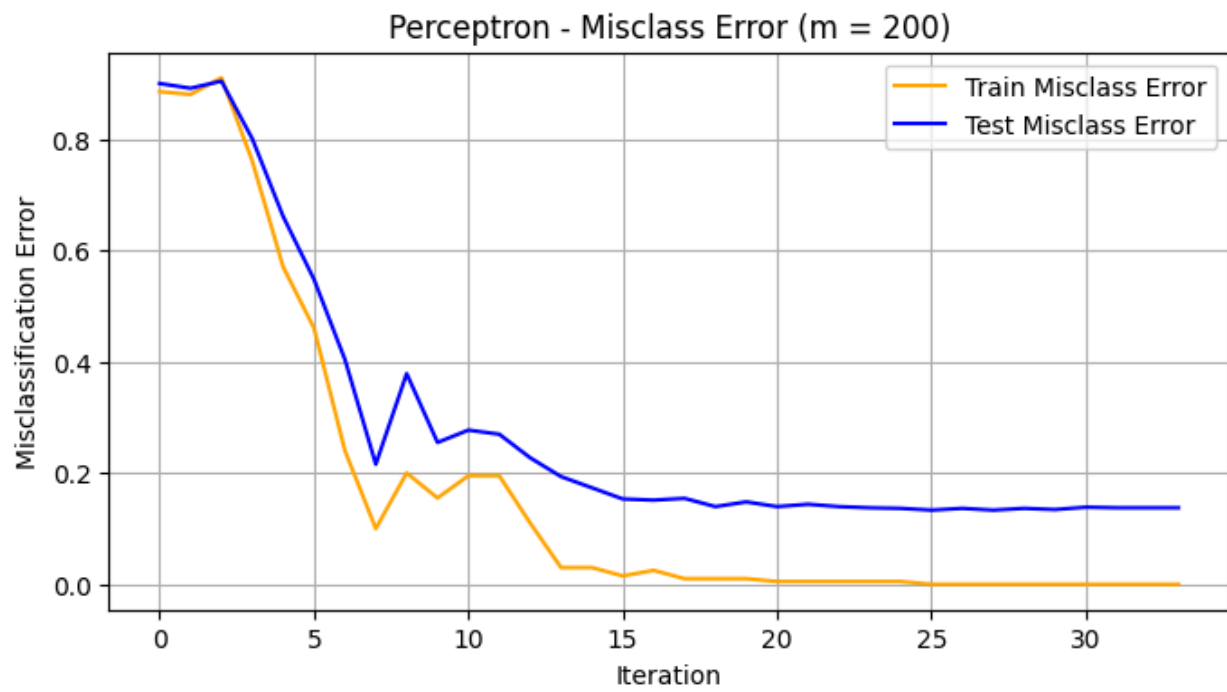
- 
- m = 200



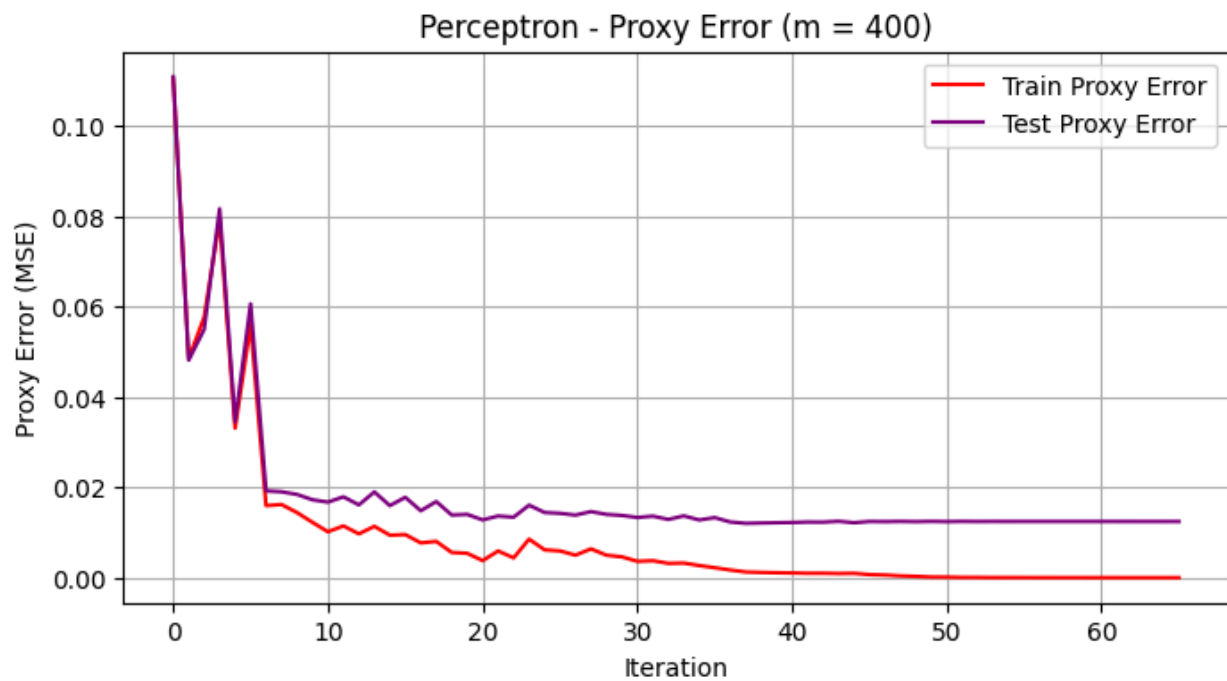


## CIS 481 Project 2: Handwritten Digits Recognition using Neural Networks

Written by: Zaynab Mourtada

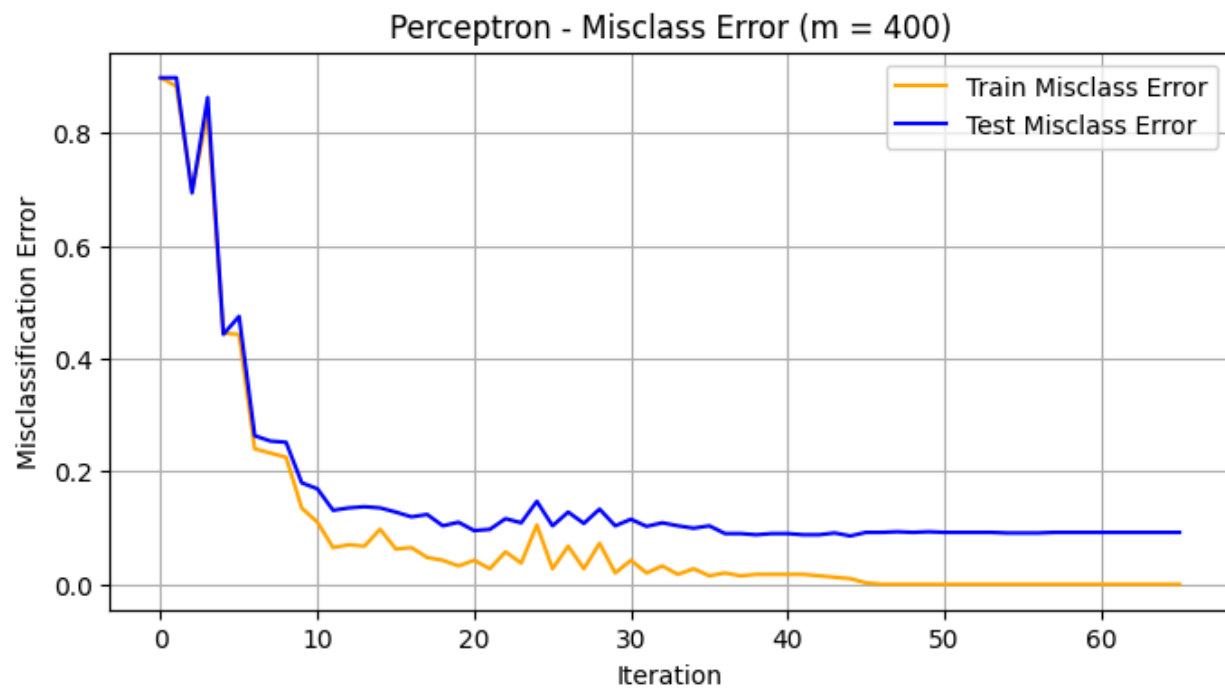


- 
- m = 400

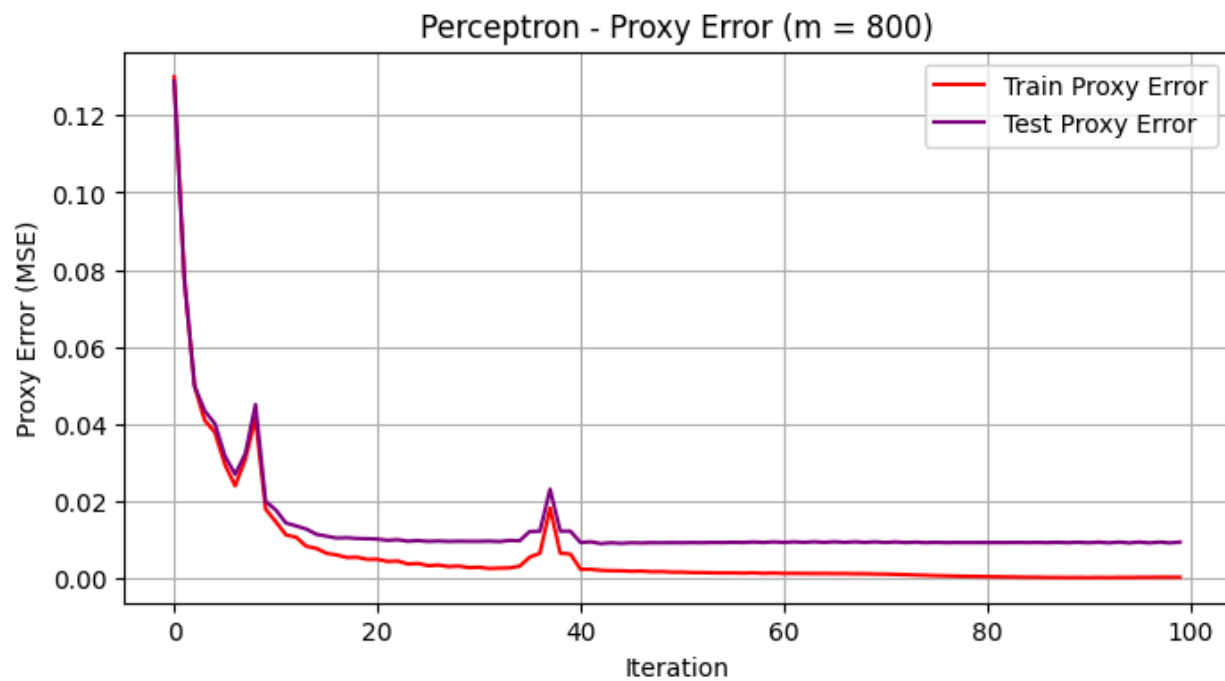


**CIS 481 Project 2: Handwritten Digits Recognition using Neural Networks**

**Written by: Zaynab Mourtada**

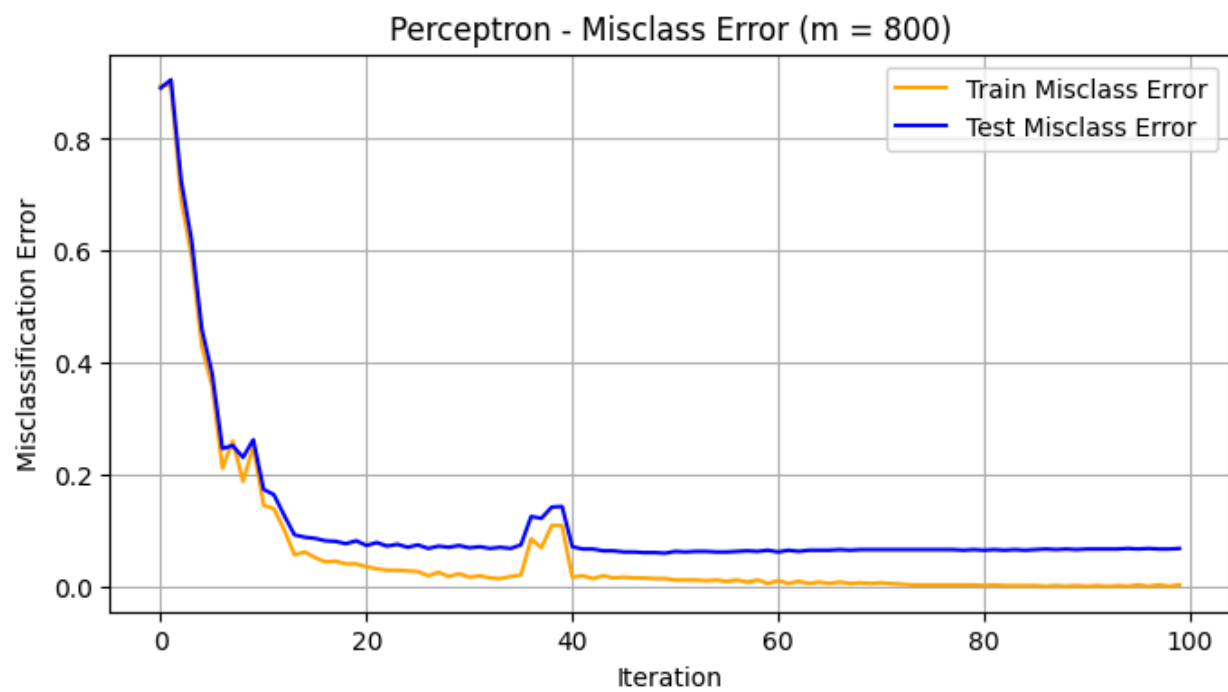


- 
- m = 800

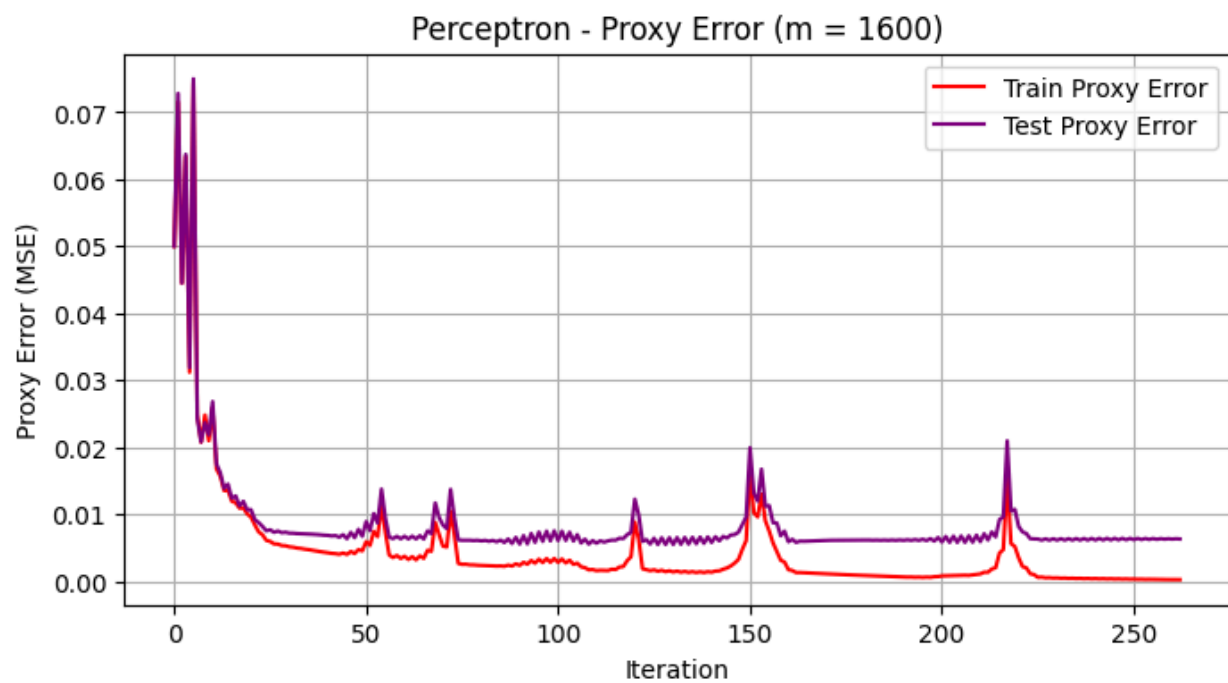


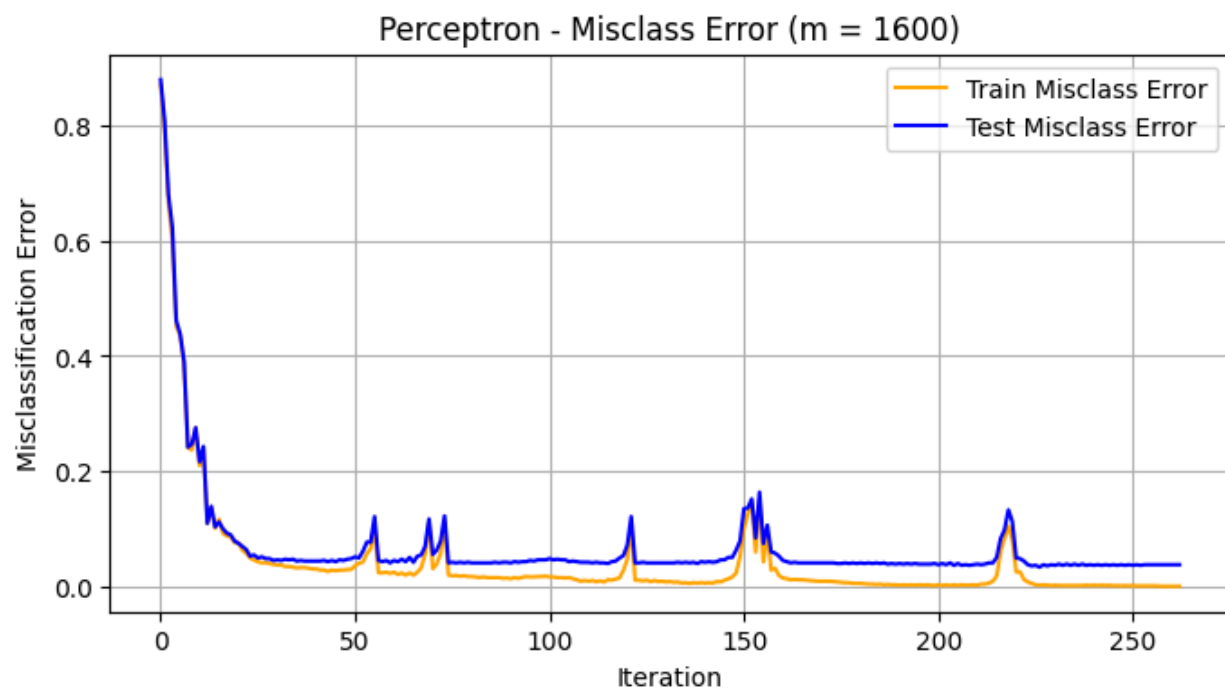
**CIS 481 Project 2: Handwritten Digits Recognition using Neural Networks**

**Written by: Zaynab Mourtada**



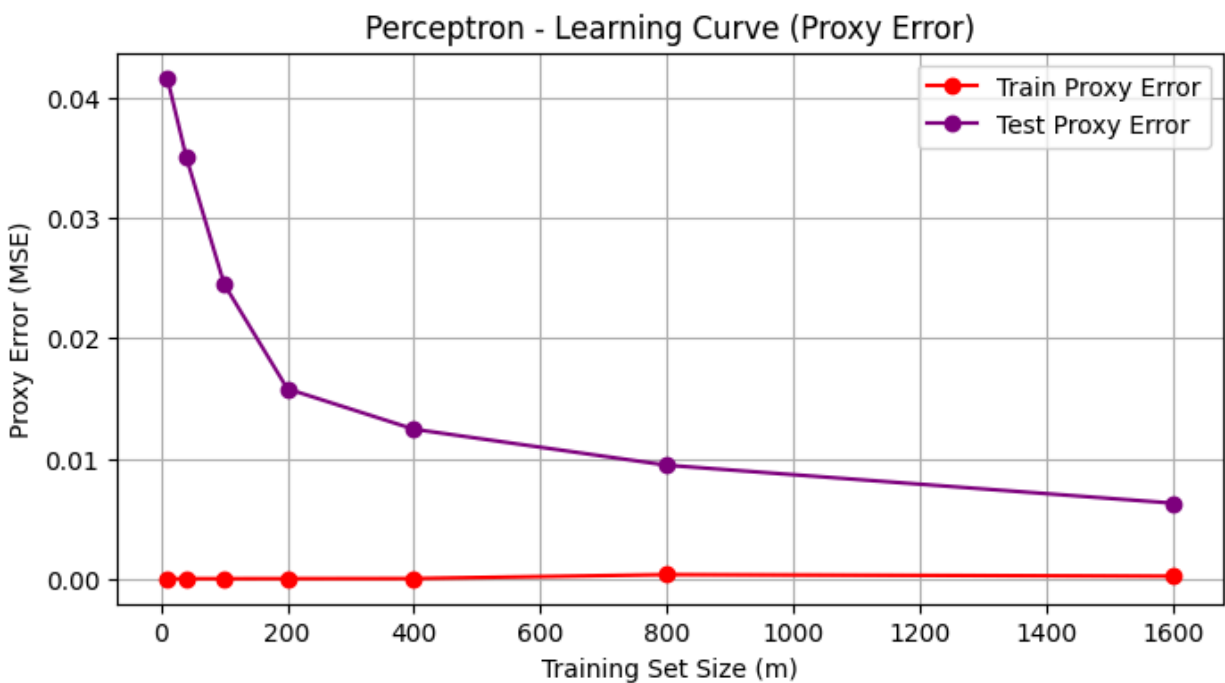
- 
- m = 1600





**b) Overall Learning Curves Across Subset Sizes**

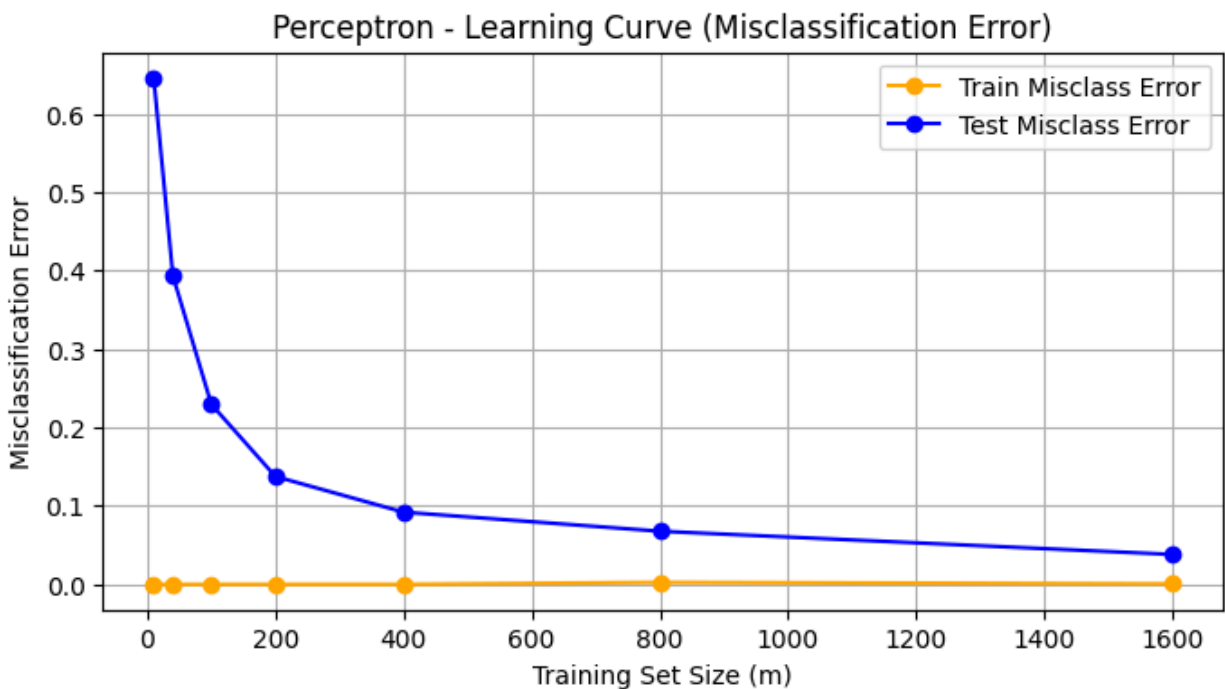
As the training set size increases, the test proxy error steadily decreases, showing improved generalization. Training proxy error stays consistently low, indicating effective learning across all subset sizes.



## ***CIS 481 Project 2: Handwritten Digits Recognition using Neural Networks***

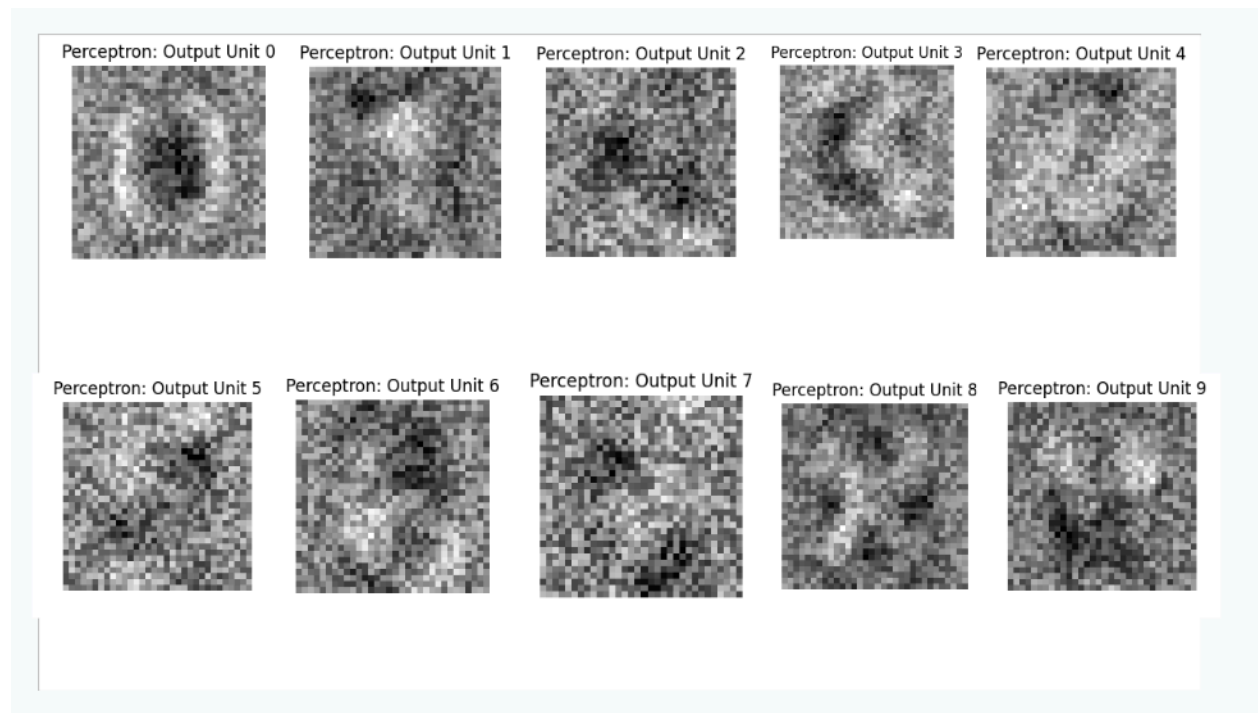
***Written by: Zaynab Mourtada***

Test misclassification error drops significantly as more training data is added, while training error remains close to zero. This suggests that larger datasets help the model generalize better and reduce overfitting.



### ***Weight Parameter Interpretation***

This section visualizes the learned weights for each output unit in the final Perceptron model. Each image shows the weight values reshaped into the original  $32 \times 32$  input format, giving a rough sense of which pixel patterns the model associates with each digit class.



## Deep Neural Network Model:

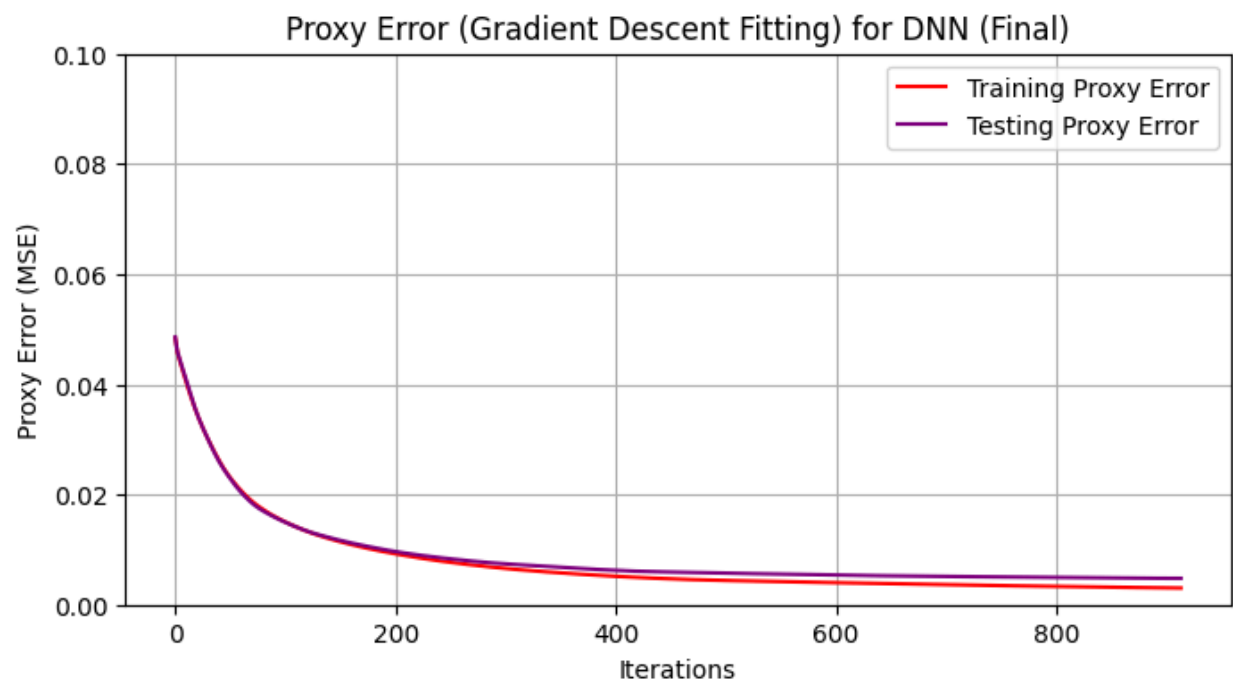
### ***Gradient Descent Fitting Curves:***

#### ***a) Training and Testing Proxy Errors:***

This graph shows how the training and testing proxy errors evolve over the course of gradient descent iterations for the final DNN model. The network had a depth of 3, which was the best-performing configuration selected during cross-validation. Both the training and testing proxy errors decrease smoothly, showing that the model is consistently minimizing error as it learns. The numbers in the screenshot are just a sample of the full output; there are many more values, but I included a portion here to highlight the general trend.

**CIS 481 Project 2: Handwritten Digits Recognition using Neural Networks**

**Written by: Zaynab Mourtada**



## *CIS 481 Project 2: Handwritten Digits Recognition using Neural Networks*

*Written by: Zaynab Mourtada*

```
>>> DNN (Final) - ALL Training Proxy Error:
[0.04840506 0.04706019 0.04584306 0.04507254 0.04436707 0.04372454
 0.04304224 0.04236933 0.04168882 0.04102417 0.04036364 0.03971888
 0.03908415 0.03846572 0.03786121 0.03727304 0.03669894 0.03613854
 0.03558998 0.03505317 0.03452842 0.03401682 0.03351899 0.0330348
 0.03256333 0.03210345 0.03165425 0.03121517 0.03078577 0.03036542
 0.02995305 0.02954721 0.02914644 0.02874982 0.02835725 0.02796917
 0.0275862 0.02720902 0.02683809 0.02647364 0.02611603 0.02576597
 0.02542436 0.02509209 0.02476979 0.02445755 0.02415513 0.0238622
 0.02357844 0.02330345 0.0230366 0.02277701 0.02252361 0.02227529
 0.02203112 0.02179055 0.02155353 0.02132048 0.02109215 0.02086922
 0.02065195 0.02044007 0.0202331 0.0200307 0.01983263 0.0196386
 0.01944821 0.01926116 0.01907766 0.01889859 0.01872509 0.01855781
 0.01839661 0.01824081 0.01808965 0.01794259 0.01779928 0.01765935
 0.01752232 0.01738764 0.01725486 0.01712374 0.0169943 0.01686672
 0.01674126 0.01661812 0.01649738 0.01637904 0.01626302 0.01614921
 0.01603752 0.01592787 0.01582017 0.01571432 0.01561019 0.01550766
 0.01540657 0.01530676 0.01520813 0.01511057 0.01501402 0.01491845
 0.01482386 0.01473023 0.01463757 0.01454586 0.01445509 0.01436525
 0.01427633 0.01418833 0.01410127 0.01401518 0.0139301 0.01384608
 0.01376318 0.01368144 0.01360091 0.0135216 0.01344353 0.01336668
 0.01329104 0.01321657 0.01314324 0.01307099 0.01299976 0.01292949
 0.01286013 0.01279161 0.0127239 0.01265696 0.01259077 0.0125253
 0.01246055 0.01239653 0.01233323 0.01227068 0.01220888 0.01214784
 0.01208756 0.01202804 0.01196927 0.01191122 0.01185387 0.01179719
```



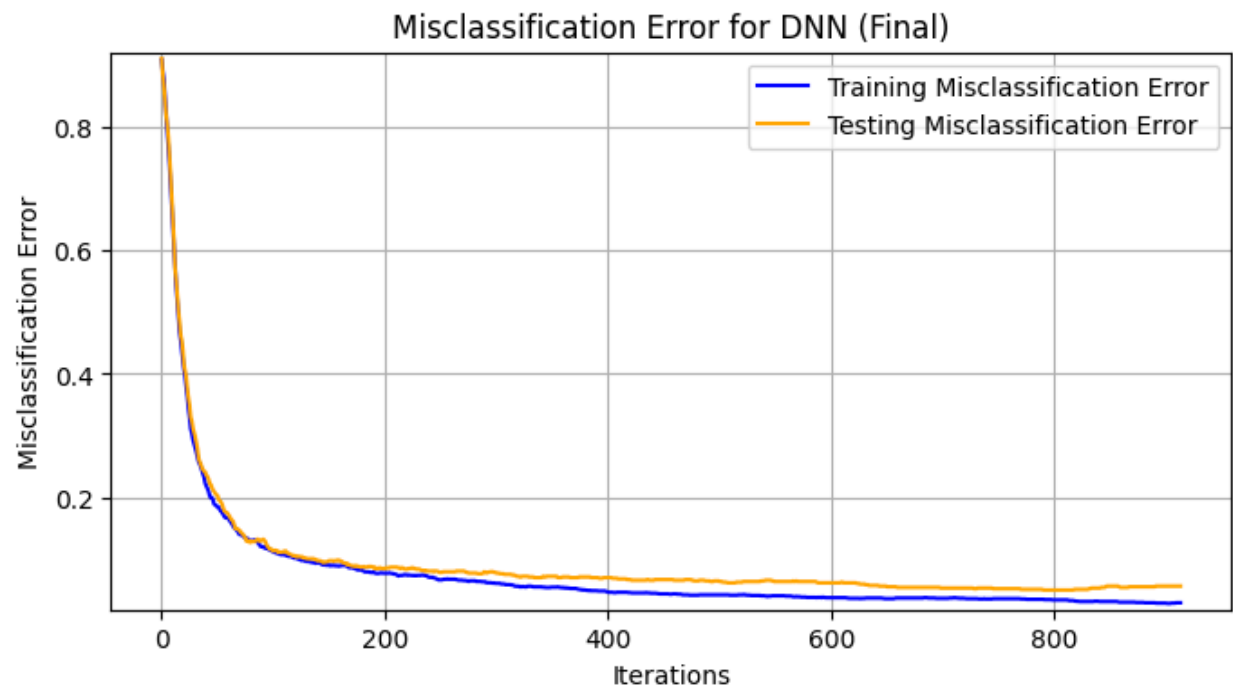
```
>>> DNN (Final) - ALL Test Proxy Error:
[0.04857088 0.04708719 0.04590314 0.0452212 0.04457048 0.04395909
0.04331695 0.04267351 0.04201957 0.04136308 0.04069662 0.04003239
0.03937243 0.03872303 0.03808306 0.03745535 0.03684114 0.03624286
0.03566194 0.03510013 0.03455836 0.03403637 0.03353314 0.03304719
0.03257622 0.03211724 0.03166721 0.03122348 0.03078413 0.03034782
0.02991359 0.02948112 0.02905123 0.02862584 0.02820733 0.02779782
0.02739869 0.0270105 0.02663319 0.02626659 0.02591078 0.02556614
0.02523275 0.02491001 0.0245968 0.02429198 0.02399466 0.02370384
0.02341835 0.02313722 0.02285993 0.02258638 0.02231666 0.02205089
0.02178929 0.02153225 0.02128037 0.02103427 0.02079446 0.02056128
0.02033493 0.02011531 0.01990195 0.01969428 0.01949194 0.01929461
0.01910197 0.01891386 0.01873054 0.01855271 0.01838114 0.01821611
0.01805733 0.0179044 0.01775721 0.01761575 0.01747971 0.01734857
0.01722182 0.01709901 0.01697967 0.01686327 0.01674926 0.01663725
0.01652701 0.01641845 0.0163115 0.01620604 0.01610187 0.01599882
0.01589672 0.01579555 0.01569536 0.0155963 0.01549852 0.01540212
0.01530712 0.01521351 0.01512121 0.01503011 0.01494013 0.01485113
0.01476304 0.01467576 0.01458925 0.0145035 0.01441851 0.01433434
0.01425104 0.01416867 0.0140873 0.01400697 0.0139277 0.01384952
0.01377247 0.01369656 0.01362183 0.01354833 0.01347609 0.01340516
0.0133356 0.01326746 0.01320077 0.01313556 0.01307183 0.01300952
0.01294857 0.01288888 0.01283031 0.01277275 0.01271605 0.01266008
0.01260473 0.01254991 0.01249553 0.01244154 0.0123879 0.01233457
0.01228155 0.01222881 0.01217635 0.01212416 0.01207222 0.01202055
0.01196912 0.01191796 0.01186705 0.01181643 0.01176609 0.01171608
0.01166642 0.01161714 0.01156827 0.01151984 0.01147188 0.01142441]
```

### ***b) Training and Testing Misclassification Errors:***

This figure displays the training and testing misclassification errors for the final DNN model with a depth of 3. The learning curve shows a sharp decrease early in training, with both curves gradually leveling off, indicating convergence. The screenshot below includes only a sample of the values plotted due to the large number of iterations.

**CIS 481 Project 2: Handwritten Digits Recognition using Neural Networks**

**Written by: Zaynab Mourtada**



## *CIS 481 Project 2: Handwritten Digits Recognition using Neural Networks*

*Written by: Zaynab Mourtada*

```
>>> DNN (Final) - ALL Training Misclassification Error:
[0.9089969 0.89038263 0.87693899 0.85005171 0.82161324 0.79420889
0.77042399 0.73991727 0.70475698 0.67166494 0.6401241 0.60599793
0.57135471 0.54291624 0.51706308 0.49534643 0.46845915 0.45501551
0.43743537 0.42244054 0.40641158 0.39555326 0.37693899 0.36297828
0.34332989 0.32936918 0.31127198 0.30403309 0.29369183 0.28748707
0.28024819 0.27352637 0.26525336 0.25801448 0.25439504 0.24974147
0.24353671 0.23733195 0.23216132 0.22285419 0.21820062 0.21458118
0.20992761 0.20268873 0.19906929 0.19958635 0.19544984 0.18872802
0.1876939 0.18665977 0.18407446 0.18407446 0.18045502 0.17735264
0.17631851 0.17269907 0.16907963 0.16701138 0.16649431 0.16649431
0.16597725 0.16184074 0.15977249 0.15667011 0.1540848 0.15046536
0.14943123 0.14684592 0.14529473 0.14115822 0.13960703 0.1401241
0.13908997 0.13702172 0.13547053 0.13391934 0.13236815 0.13288521
0.13133402 0.12978283 0.12926577 0.12926577 0.1302999 0.13081696
0.12926577 0.12926577 0.12668046 0.12512927 0.12150982 0.11995863
0.11995863 0.11892451 0.11995863 0.11737332 0.11685626 0.11633919
0.11685626 0.11427094 0.11375388 0.11271975 0.11168563 0.11116856
0.11013444 0.10961737 0.10910031 0.10806618 0.10754912 0.10703206
0.10703206 0.10806618 0.10703206 0.10599793 0.10599793 0.10444674
0.10444674 0.10444674 0.10392968 0.10289555 0.10237849 0.10134436
0.10134436 0.10134436 0.10031024 0.09927611 0.09927611 0.09772492
0.09824199 0.09875905 0.09720786 0.09617373 0.09617373 0.09617373
0.09565667 0.09565667 0.09462254 0.09462254 0.09410548 0.09410548
0.09307135 0.09410548 0.09358842 0.09358842 0.09307135 0.09203723
0.09048604 0.09048604 0.09048604 0.08996898 0.08996898 0.08996898
0.08996898 0.08893485 0.08945191 0.08945191 0.08945191 0.08945191
0.08893485 0.08893485 0.08893485 0.08841779 0.08893485 0.08945191
0.08945191 0.08996898 0.08945191 0.08945191 0.08893485 0.0868666]
```

```
>>> DNN (Final) - ALL Test Misclassification Error:
[0.91120507 0.88477801 0.86786469 0.84143763 0.8192389 0.80443975
 0.7832981 0.75475687 0.7230444 0.68816068 0.65221987 0.61627907
 0.57082452 0.55179704 0.52536998 0.50211416 0.48520085 0.46300211
 0.45454545 0.43763214 0.41437632 0.40486258 0.39429175 0.37526427
 0.36046512 0.34883721 0.32980973 0.32346723 0.31183932 0.30549683
 0.29809725 0.28858351 0.27906977 0.26427061 0.25792812 0.25369979
 0.24841438 0.24418605 0.24207188 0.2410148 0.23572939 0.23361522
 0.22938689 0.22198732 0.22093023 0.21564482 0.21141649 0.20930233
 0.20613108 0.20401691 0.20295983 0.19661734 0.19450317 0.19238901
 0.18921776 0.18287526 0.17758985 0.17653277 0.17547569 0.1744186
 0.17124736 0.16807611 0.16490486 0.1627907 0.16067653 0.15539112
 0.14904863 0.14904863 0.14904863 0.14693446 0.14587738 0.14270613
 0.14164905 0.13953488 0.13742072 0.13319239 0.12896406 0.12896406
 0.13002114 0.12684989 0.12896406 0.12790698 0.12790698 0.12896406
 0.12896406 0.13002114 0.13213531 0.13002114 0.12896406 0.13002114
 0.13002114 0.13213531 0.13002114 0.12790698 0.12262156 0.1205074
 0.11733615 0.11627907 0.11627907 0.1141649 0.11310782 0.1141649
 0.1141649 0.11310782 0.11099366 0.11205074 0.11099366 0.11099366
 0.11099366 0.11205074 0.11205074 0.11310782 0.11099366 0.10887949
 0.10887949 0.10782241 0.10676533 0.10570825 0.10570825 0.10570825
 0.10359408 0.10359408 0.10465116 0.10359408 0.10359408 0.10359408
 0.10359408 0.10147992 0.10147992 0.10147992 0.10042283 0.10042283
 0.10042283 0.10042283 0.10042283 0.10042283 0.10042283 0.09936575
 0.09936575 0.09725159 0.09725159 0.09725159 0.09725159 0.0961945
 0.09513742 0.09408034 0.09513742 0.09513742 0.09513742 0.0961945
 0.09725159 0.09725159 0.09725159 0.09725159 0.0961945 0.09725159
 0.09725159 0.09725159 0.09830867 0.09725159 0.09725159 0.0961945
 0.09408034 0.09408034 0.09302326 0.09302326 0.09090909 0.09196617
 0.08985201 0.08985201 0.08879493 0.08879493 0.08879493 0.08773784
 0.08879493 0.08879493 0.08773784 0.08773784 0.08668076 0.08668076]
```

### *c) Evaluation of Trial Data Set:*

The DNN misclassified one digit on the trial set, resulting in a 10% error. This might be due to slight overfitting or unnecessary complexity for such a small set of examples. Although the model is more powerful, its added depth may have made it more sensitive to minor variations in the data.

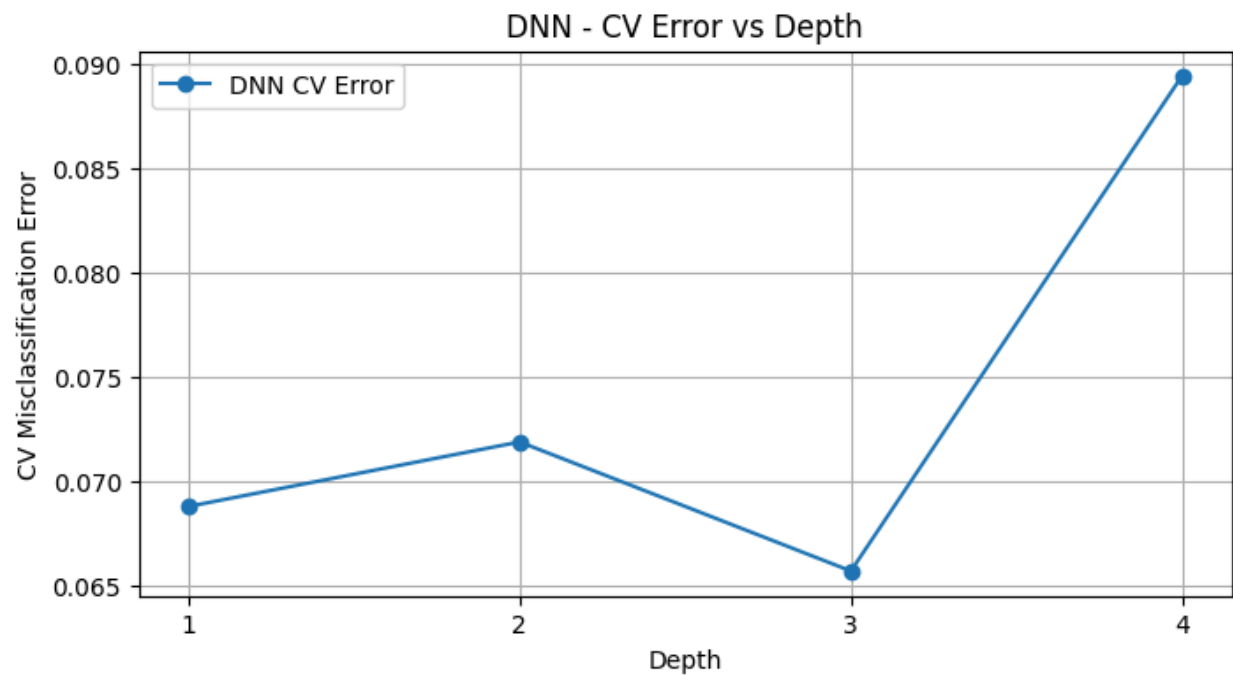
### *CV Misclassification Errors:*

Based on 3-fold cross-validation, the best-performing depth for the DNN model was determined to be 3.

## ***CIS 481 Project 2: Handwritten Digits Recognition using Neural Networks***

***Written by: Zaynab Mourtada***

Depth	Average CV Error
1	0.0688
2	0.0719
3	0.0657
4	0.0895



### ***Learning Curves:***

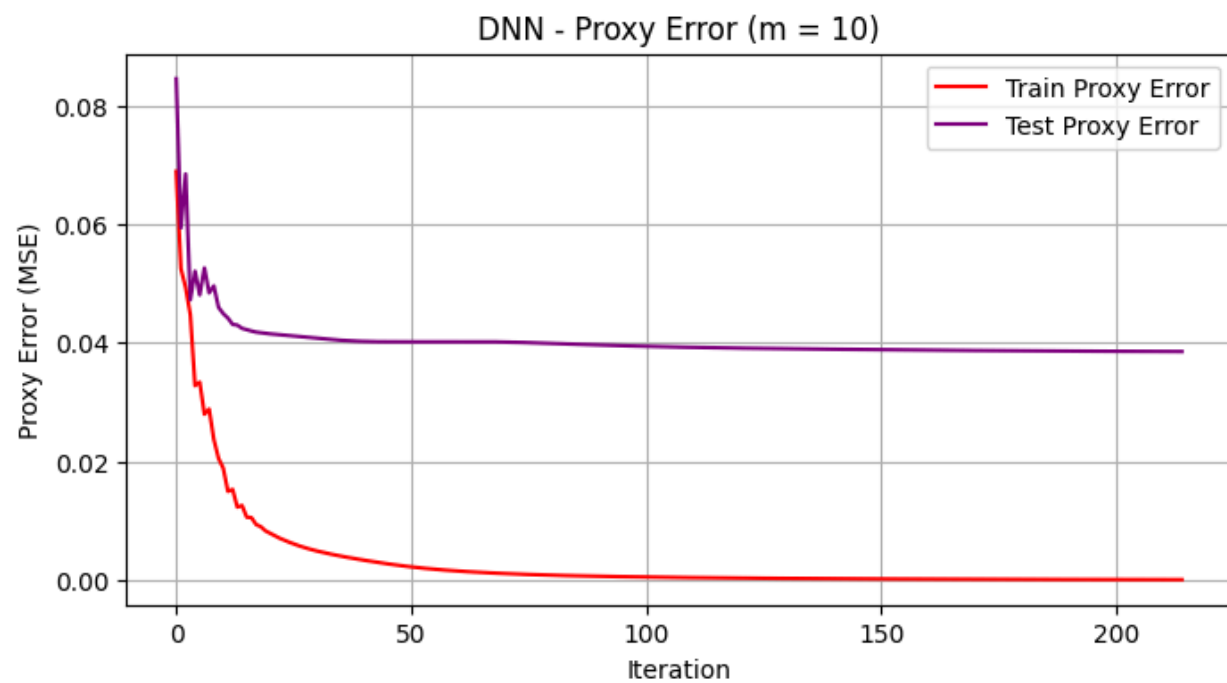
#### ***a) Learning Curves by Subset Size (Proxy & Misclassification Error vs Iterations)***

These graphs show how the training and testing errors change across iterations for different subset sizes, capturing how the DNN model learns with more data over time.

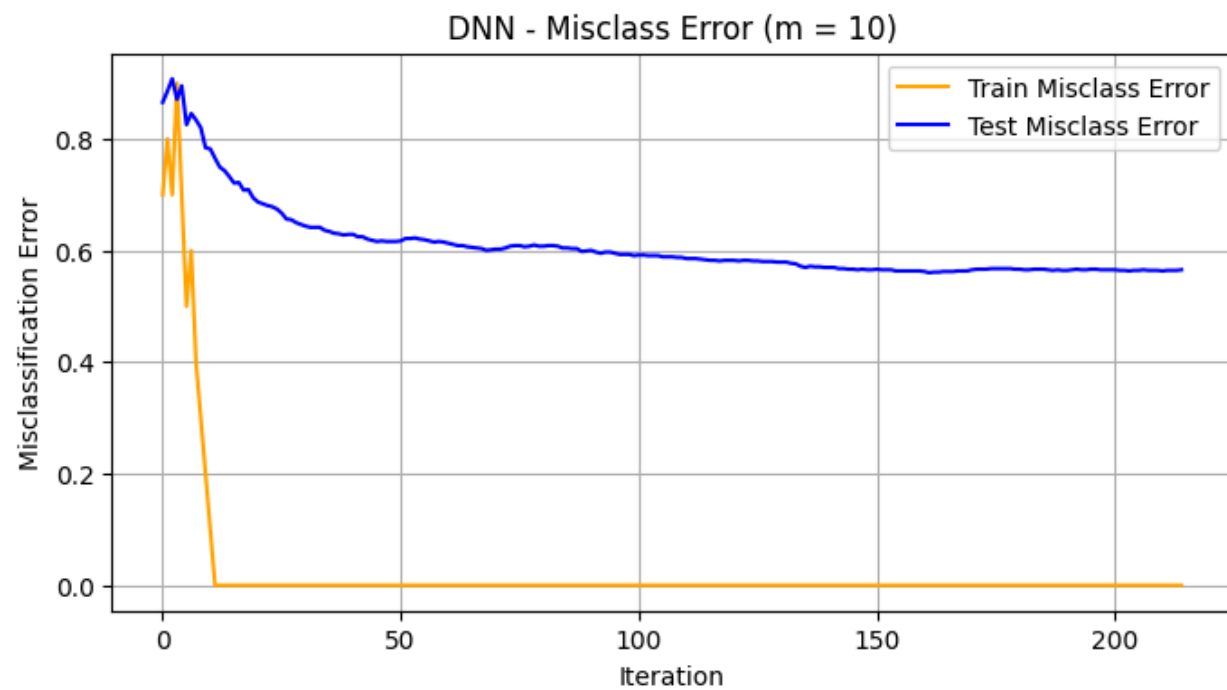
- $m = 10$

**CIS 481 Project 2: Handwritten Digits Recognition using Neural Networks**

**Written by: Zaynab Mourtada**



○



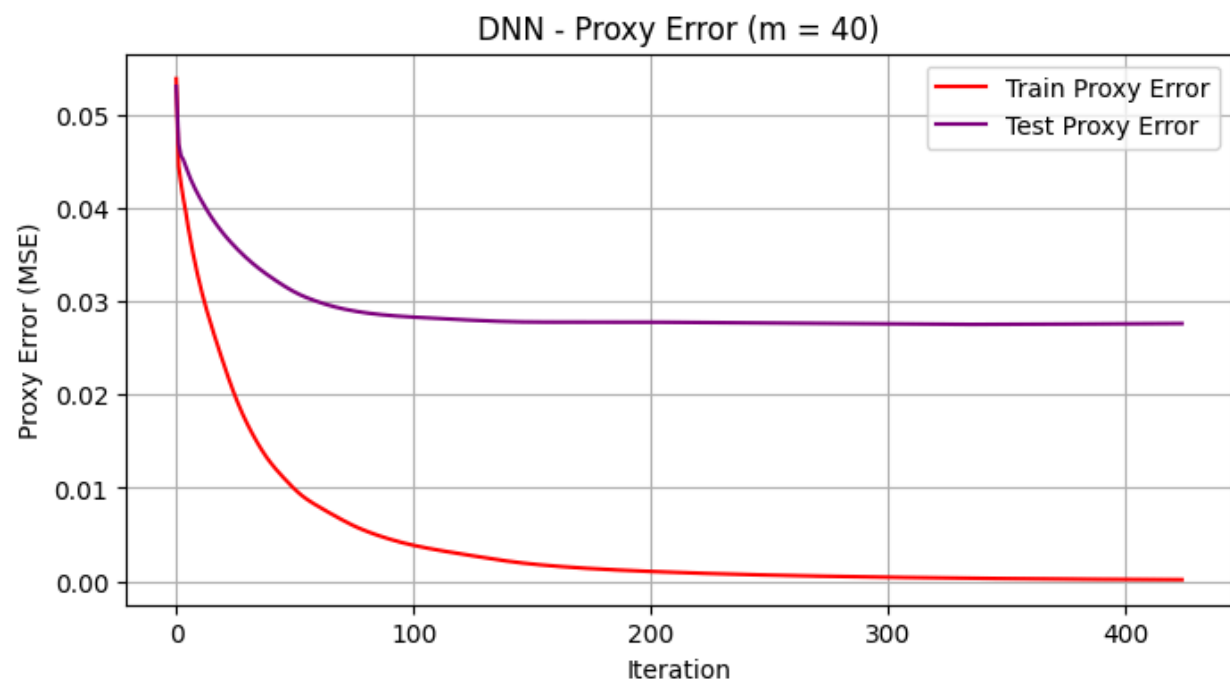
○

- m = 40

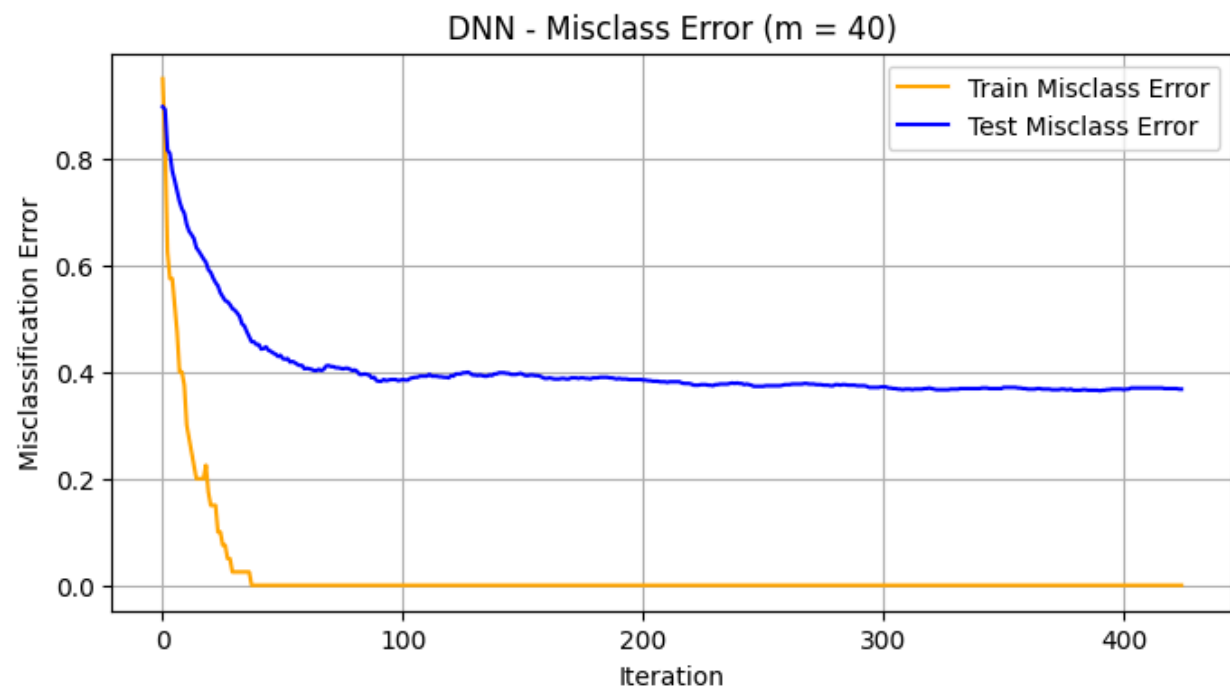


## CIS 481 Project 2: Handwritten Digits Recognition using Neural Networks

Written by: Zaynab Mourtada



○

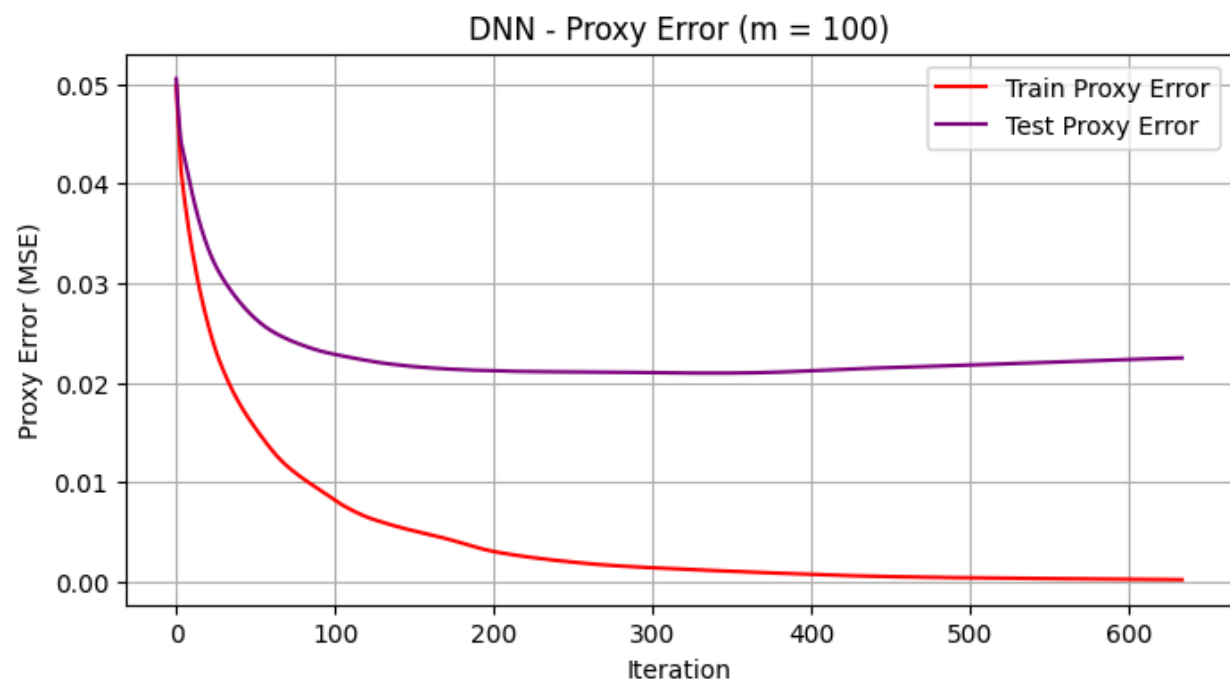


○

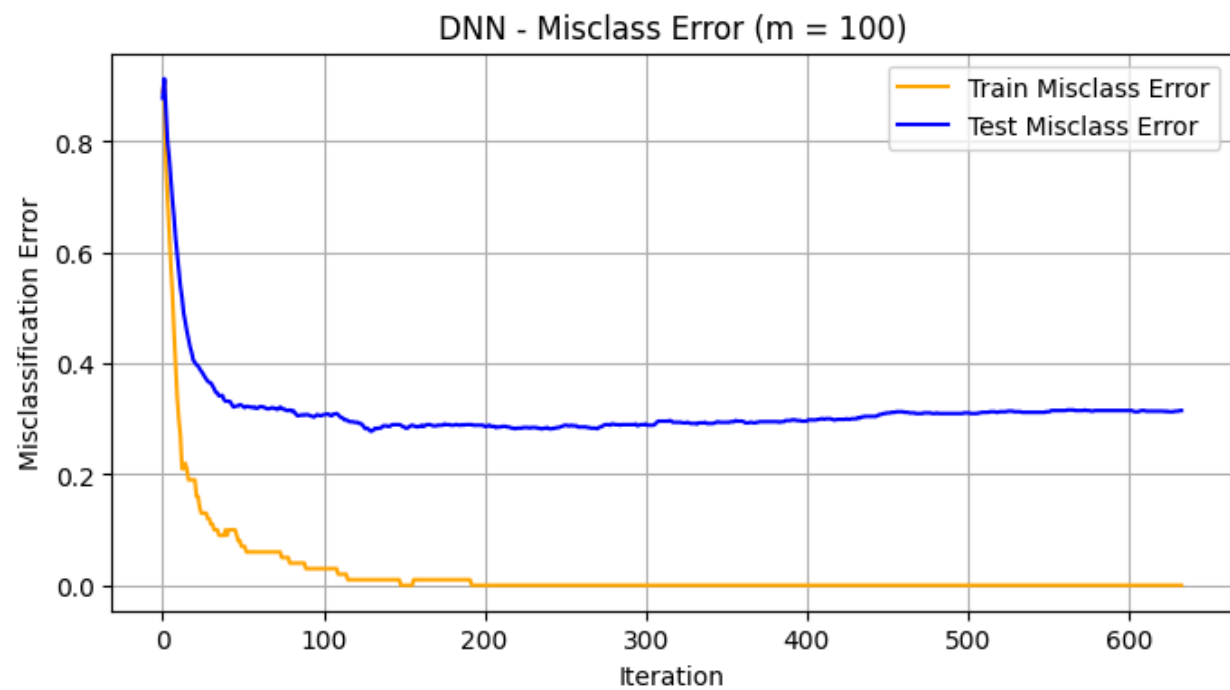
- m = 100

## CIS 481 Project 2: Handwritten Digits Recognition using Neural Networks

Written by: Zaynab Mourtada



○



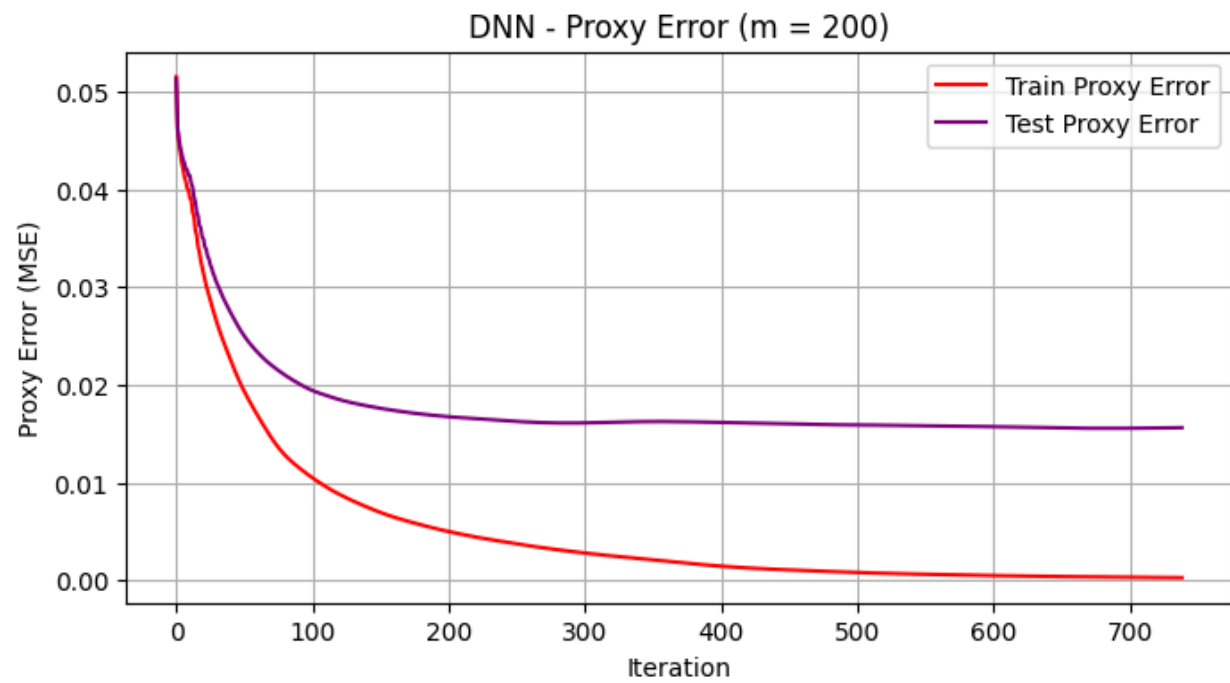
○

- m = 200

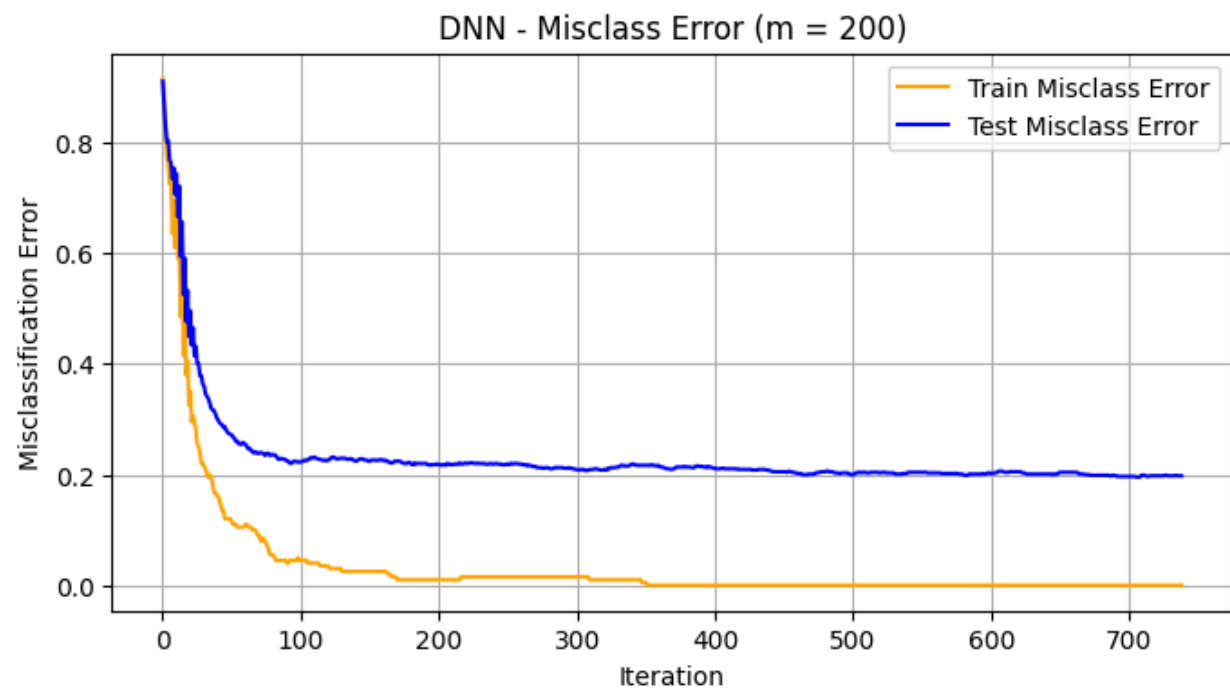


**CIS 481 Project 2: Handwritten Digits Recognition using Neural Networks**

**Written by: Zaynab Mourtada**



○

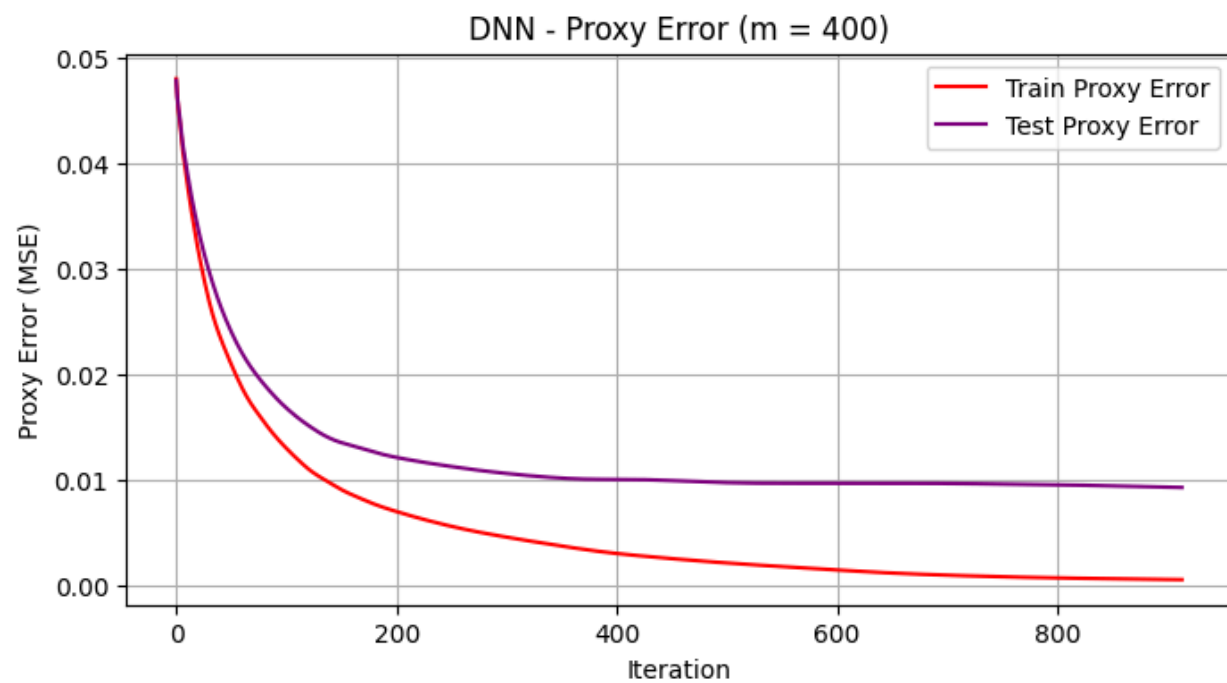


○

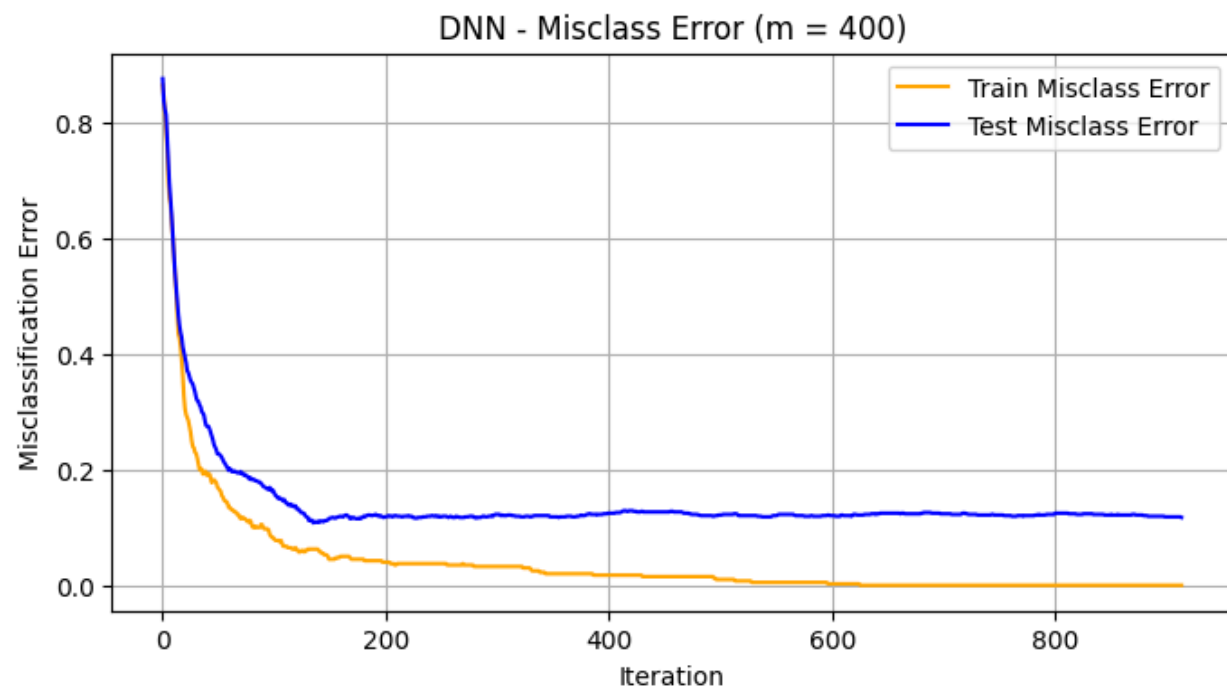
- m = 400

**CIS 481 Project 2: Handwritten Digits Recognition using Neural Networks**

**Written by: Zaynab Mourtada**



○

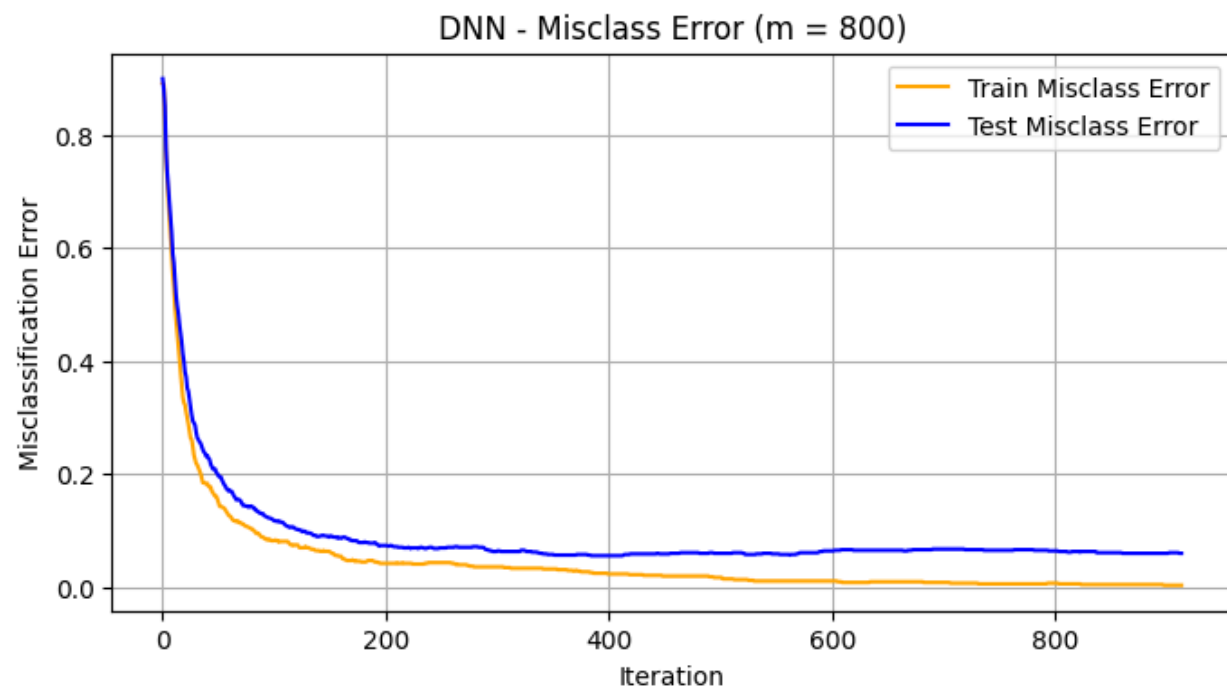
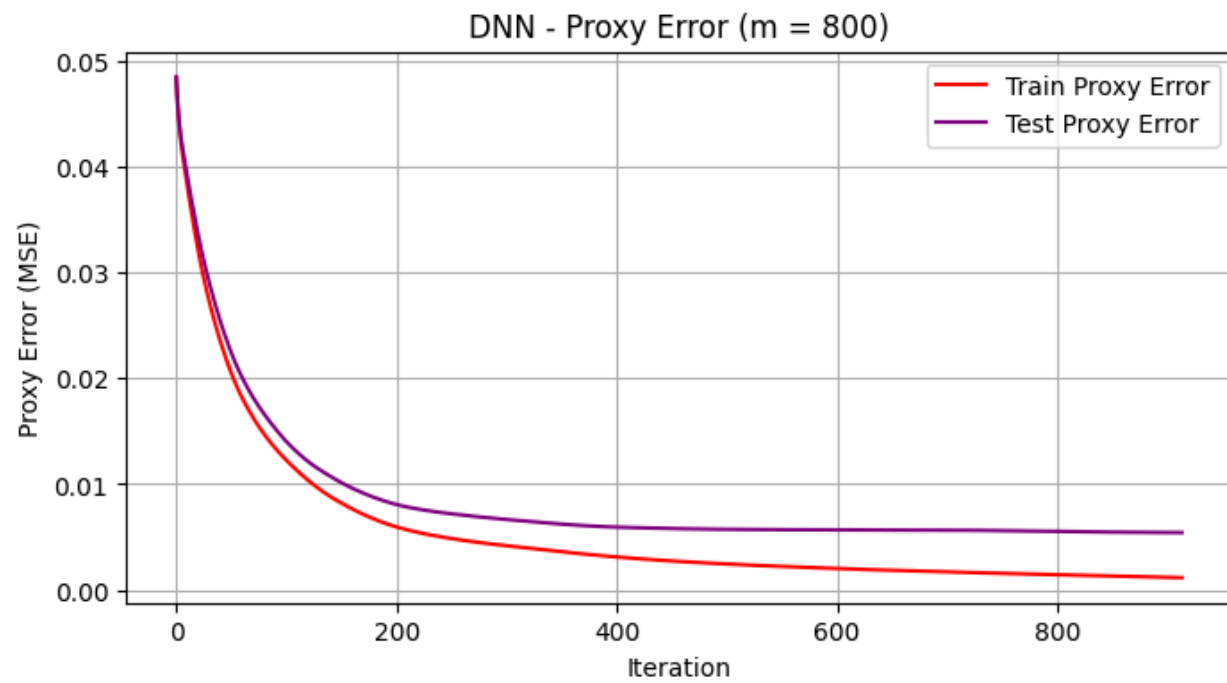


○

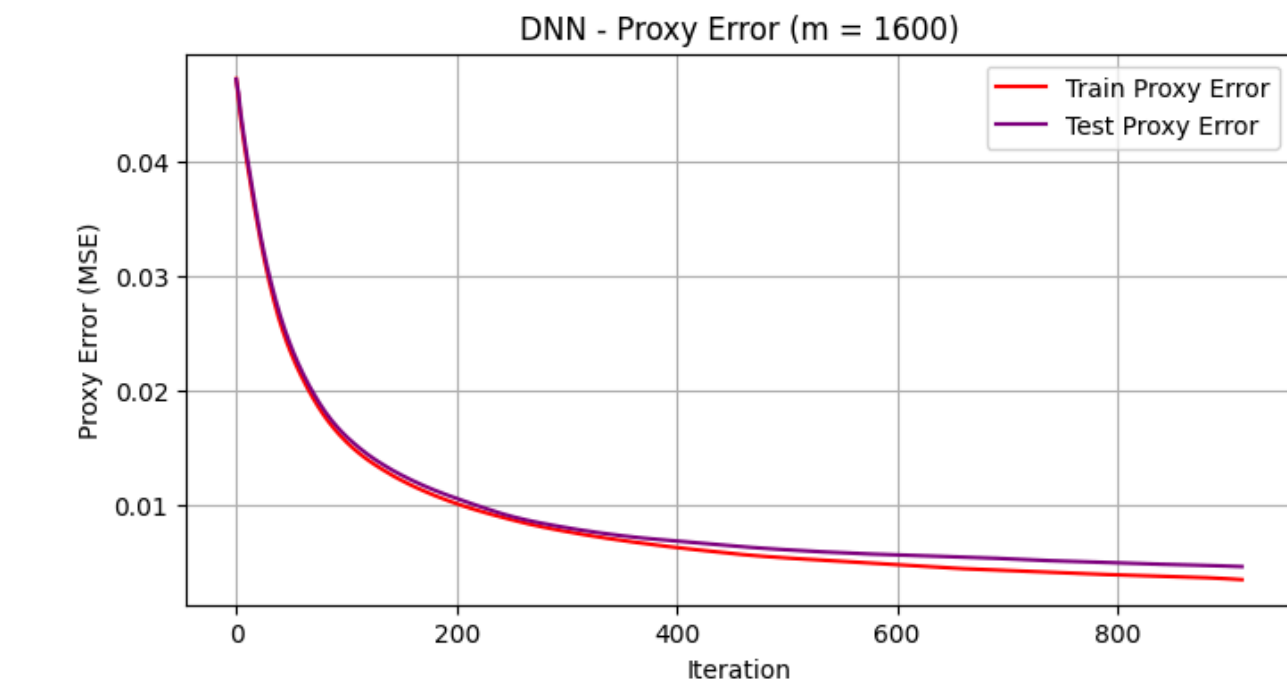
- m = 800

## CIS 481 Project 2: Handwritten Digits Recognition using Neural Networks

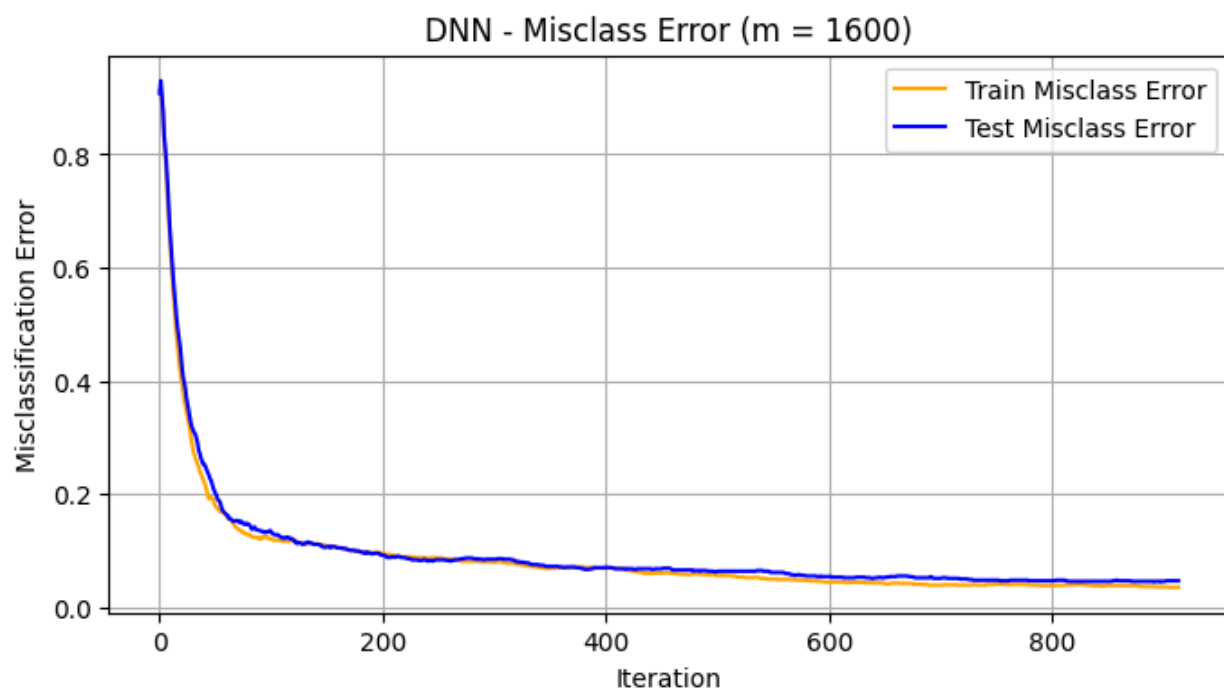
Written by: Zaynab Mourtada



- m = 1600



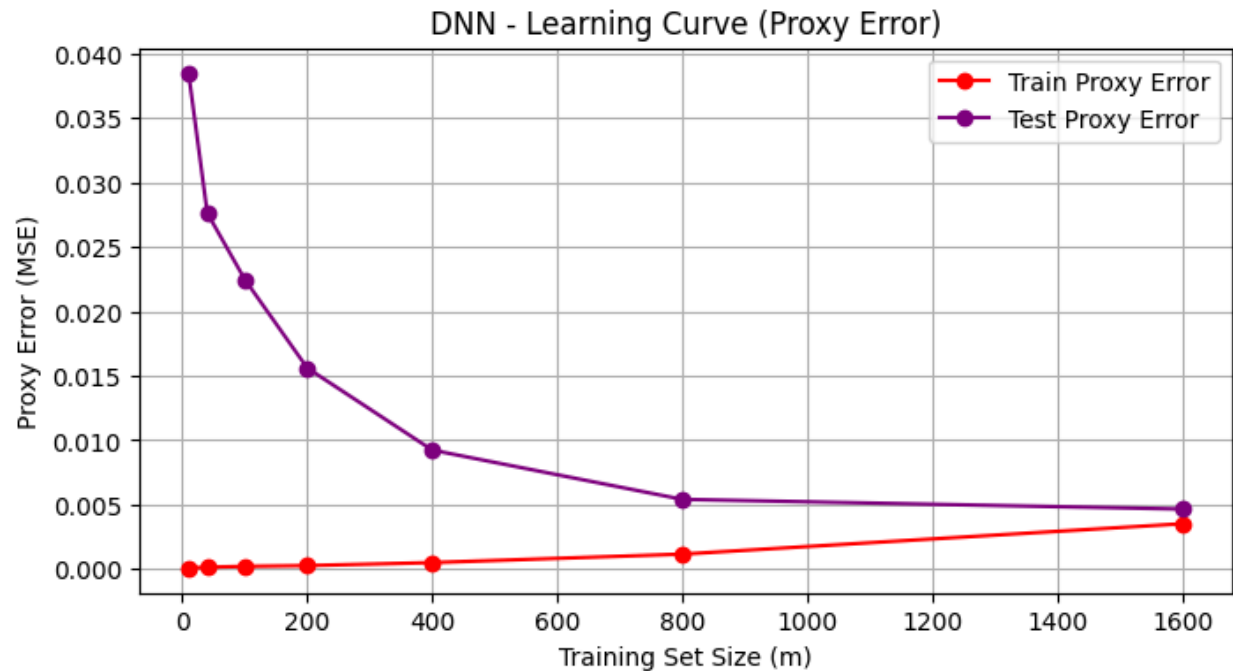
○



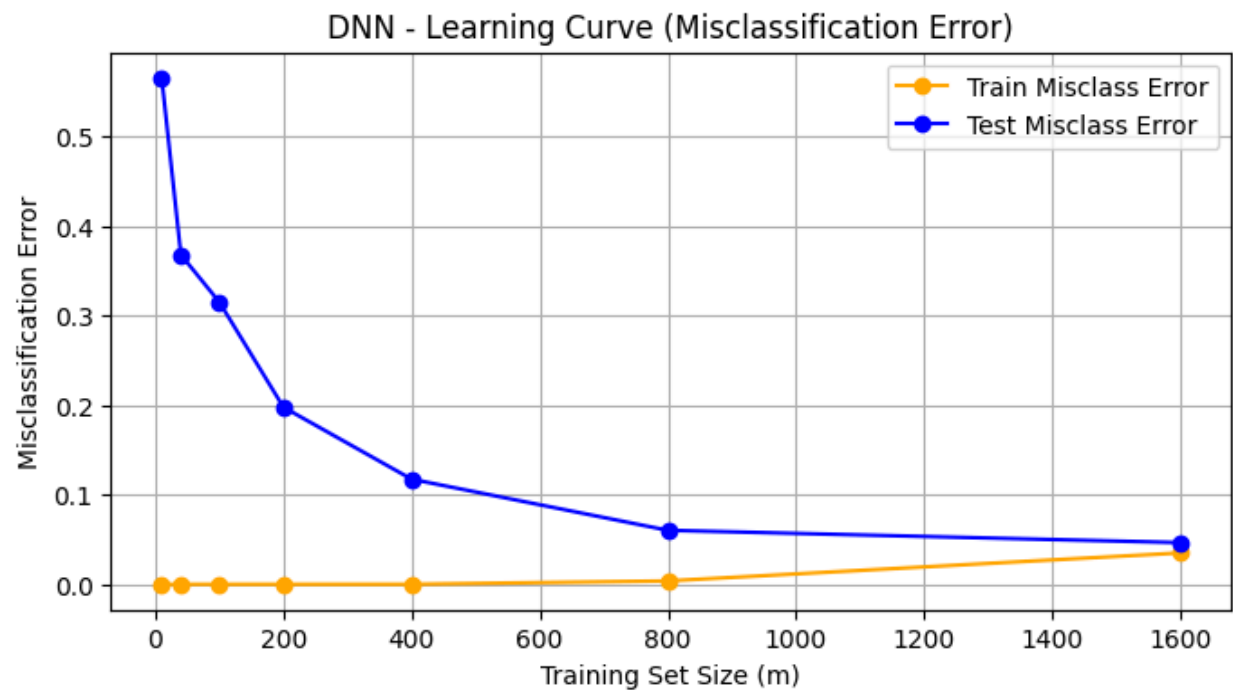
○

### ***b) Overall Learning Curves Across Subset Sizes***

As the training set size increases, the test proxy error steadily decreases, indicating that the DNN generalizes better with more data. The training proxy error stays low throughout, showing the model fits the training data well even at small subset sizes.



The test misclassification error drops sharply as the training set size increases, reflecting improved performance on unseen data. Training misclassification error remains near zero, suggesting strong memorization of the training data regardless of size.

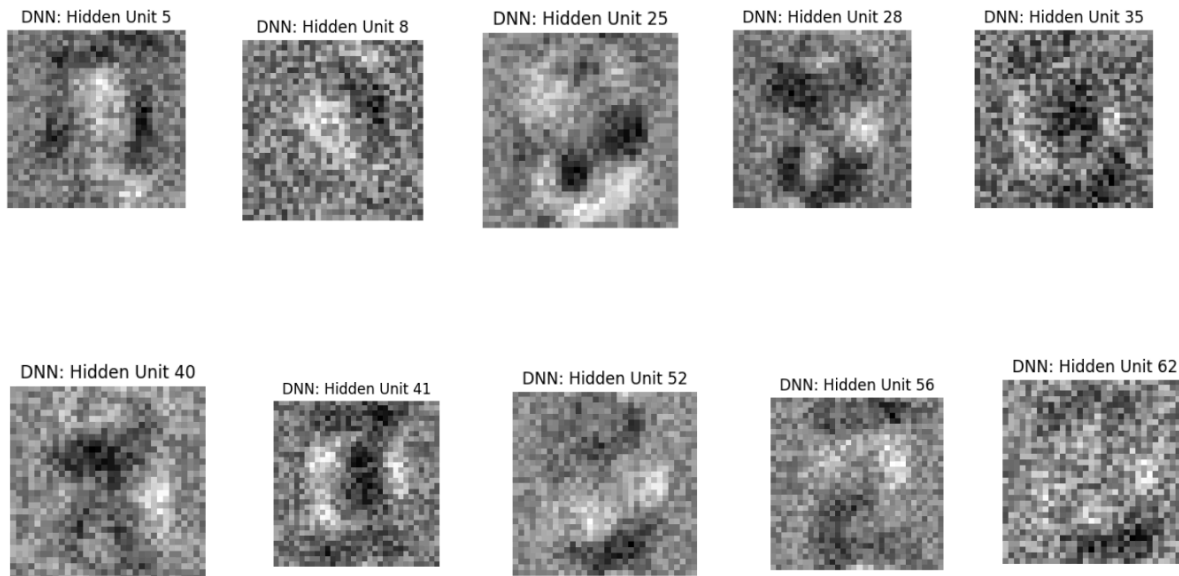


## ***Weight Parameter Interpretation***

## ***CIS 481 Project 2: Handwritten Digits Recognition using Neural Networks***

***Written by: Zaynab Mourtada***

These visualizations show the weight patterns for 10 randomly selected hidden units from the first hidden layer of the DNN. Each image reflects how that unit responds to different input pixels. Some weights resemble digit-like patterns or strokes, suggesting that the network is learning useful internal features from the data.



## **Two-Layer Deep Neural Network Model:**

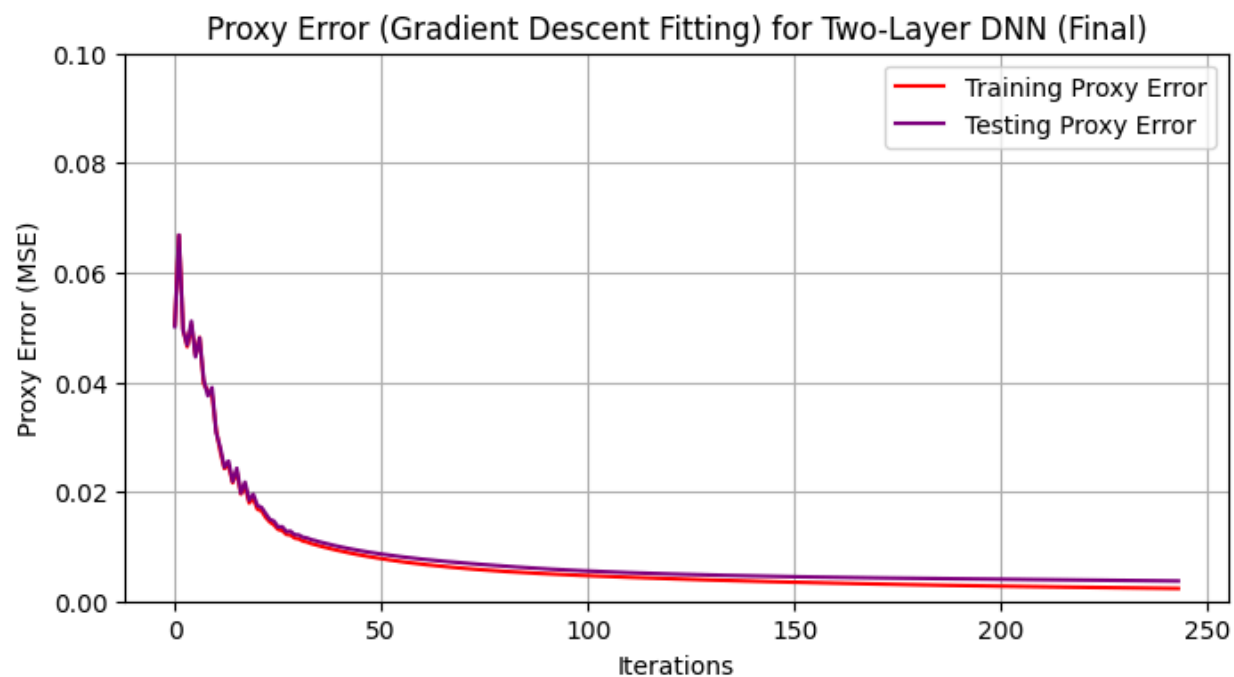
### ***Gradient Descent Fitting Curves:***

#### ***a) Training and Testing Proxy Errors:***

This graph tracks how the training and testing proxy errors (MSE) changed over time as the two-layer deep neural network trained using gradient descent. Early on, the error dropped quickly, which shows that the model was learning fast. Eventually, both curves leveled off, suggesting it had reached convergence. The learning rate used was 4, which I picked based on cross-validation results. Additionally, the screenshot below only shows a sample of the full output since there were too many values to fit in one image.

## CIS 481 Project 2: Handwritten Digits Recognition using Neural Networks

Written by: Zaynab Mourtada



```
>>> Two-Layer DNN (Final) - ALL Training Proxy Error:
[0.05056903 0.06678696 0.04964227 0.04644832 0.0504587 0.04494504
 0.04811435 0.03970806 0.03838668 0.03771146 0.03124569 0.02735614
 0.02423114 0.02480629 0.02161911 0.02362253 0.01955826 0.02101162
 0.01792021 0.01876625 0.01678215 0.01650708 0.01529979 0.0144663
 0.01394401 0.0130972 0.01291635 0.01219384 0.01208834 0.01152945
 0.01139585 0.01099315 0.01081953 0.01052824 0.01033579 0.01010669
 0.00991418 0.0097166 0.00953379 0.0093548 0.00918417 0.00901883
 0.00886022 0.00870719 0.00856017 0.0084181 0.0082808 0.00814749
 0.00801796 0.00789189 0.00776926 0.00765017 0.00753487 0.00742354
 0.00731631 0.00721308 0.00711369 0.00701786 0.00692529 0.00683569
 0.0067488 0.00666446 0.00658255 0.00650301 0.00642576 0.00635065
 0.00627752 0.00620624 0.00613674 0.00606897 0.00600293 0.00593864
 0.00587609 0.00581526 0.00575616 0.00569879 0.00564314 0.00558921
 0.00553693 0.00548622 0.00543699 0.00538913 0.00534256 0.00529723
 0.00525307 0.00521004 0.00516808 0.0051271 0.00508705 0.00504786
 0.00500947 0.00497184 0.00493495 0.00489875 0.00486324 0.00482841
 0.00479422 0.00476066 0.00472772 0.00469536 0.00466358 0.00463234
 0.00460163 0.00457143 0.00454169 0.00451239 0.0044835 0.00445498
 0.0044268 0.00439894 0.00437138 0.00434408 0.00431703 0.00429022
 0.00426364 0.00423728 0.00421113 0.0041852 0.00415949 0.00413399
 0.00410873 0.00408369 0.00405889 0.00403434 0.00401005 0.00398601
 0.00396223 0.00393873 0.00391549 0.00389253 0.00386983 0.00384741]
```

```
>>> Two-Layer DNN (Final) - ALL Test Proxy Error:
[0.05011491 0.06685365 0.04950932 0.04677196 0.05111535 0.0446068
0.04807291 0.04039291 0.03749455 0.03901167 0.03081945 0.02808457
0.02446169 0.02560038 0.02183912 0.02428652 0.01975381 0.02174422
0.0184325 0.01953001 0.01742687 0.0171248 0.01598896 0.01493728
0.01462595 0.01354971 0.01360498 0.01272604 0.01278456 0.0121512
0.01210668 0.01167581 0.0115449 0.01124514 0.01106863 0.01084051
0.01065204 0.01045944 0.01027797 0.01010502 0.00993732 0.00977904
0.00962588 0.00948054 0.00934038 0.00920651 0.00907761 0.00895406
0.00883515 0.00872077 0.00861028 0.0085033 0.00839938 0.0082983
0.00819996 0.0081044 0.00801169 0.00792192 0.0078351 0.00775123
0.00767022 0.00759192 0.00751611 0.00744254 0.00737094 0.00730104
0.00723265 0.00716562 0.0070999 0.00703548 0.00697236 0.00691048
0.00684972 0.00678991 0.0067309 0.00667264 0.0066151 0.00655835
0.00650244 0.00644744 0.00639331 0.00634 0.00628746 0.00623562
0.0061845 0.00613414 0.00608467 0.0060362 0.00598886 0.00594274
0.00589792 0.00585441 0.00581218 0.0057712 0.00573141 0.00569274
0.00565515 0.00561857 0.00558294 0.00554822 0.00551436 0.00548131
0.00544903 0.0054175 0.00538668 0.00535654 0.00532704 0.00529819
0.00526994 0.00524229 0.00521523 0.00518876 0.00516285 0.00513752
0.00511275 0.00508853 0.00506486 0.00504172 0.00501909 0.00499695
0.0049753 0.00495409 0.00493333 0.00491298 0.00489303 0.00487346
```

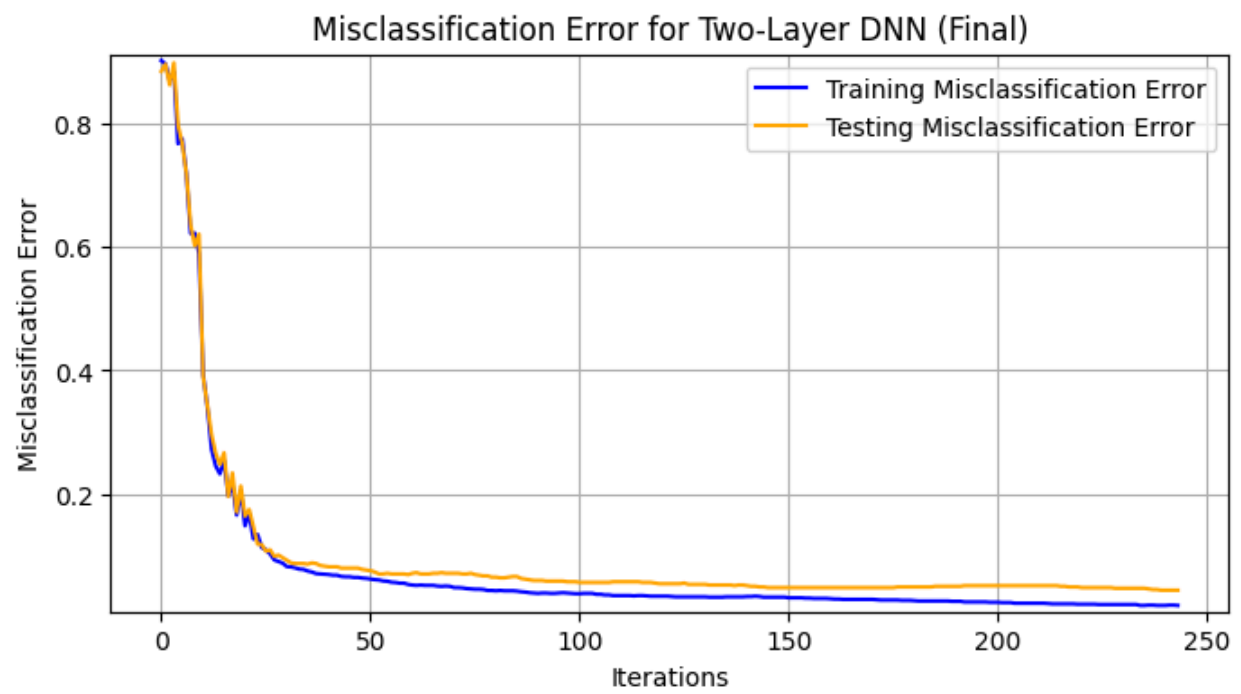
### ***b) Training and Testing Misclassification Errors:***

This graph shows how the training and testing misclassification errors changed during gradient descent for the final two-layer DNN. Right from the start, both errors dropped significantly, which indicates that the model was quickly improving. Over time, they flattened out and stayed close together, which is usually a good sign that the model is generalizing well. The learning rate used was 4; the best one found from cross-validation. And just like before, the screenshot only includes a portion of the full output values; there were too many to show all at once.



**CIS 481 Project 2: Handwritten Digits Recognition using Neural Networks**

**Written by: Zaynab Mourtada**



## *CIS 481 Project 2: Handwritten Digits Recognition using Neural Networks*

*Written by: Zaynab Mourtada*

```
>>> Two-Layer DNN (Final) - ALL Training Misclassification Error:
[0.9033092 0.89762151 0.87125129 0.885212 0.76835574 0.77611169
0.71871768 0.62254395 0.62306101 0.58996898 0.39348501 0.3417787
0.27249224 0.2450879 0.23216132 0.25387797 0.19648397 0.21561531
0.16546019 0.20423992 0.14788004 0.16752844 0.12668046 0.13391934
0.11271975 0.10910031 0.10237849 0.09255429 0.08996898 0.08738366
0.08169597 0.0811789 0.07962771 0.07755946 0.0770424 0.07445708
0.07290589 0.07032058 0.06980352 0.06928645 0.06876939 0.06773526
0.06773526 0.06566701 0.06514995 0.06514995 0.06411582 0.06411582
0.0630817 0.06256463 0.06153051 0.06049638 0.05997932 0.05842813
0.05791107 0.05635988 0.05584281 0.05480869 0.05480869 0.05274043
0.05170631 0.05118925 0.05170631 0.05118925 0.05118925 0.05118925
0.05015512 0.04963806 0.05015512 0.04963806 0.0475698 0.0475698
0.04705274 0.04601861 0.04550155 0.04498449 0.04498449 0.04446743
0.04291624 0.04291624 0.04239917 0.04291624 0.04239917 0.04239917
0.04239917 0.04188211 0.04084798 0.04033092 0.03929679 0.03877973
0.03826267 0.03877973 0.03877973 0.03826267 0.03826267 0.03877973
0.03929679 0.03877973 0.03826267 0.0377456 0.0377456 0.03826267
0.03826267 0.03826267 0.03671148 0.03671148 0.03567735 0.03567735
0.03464323 0.03464323 0.03464323 0.03464323 0.03464323 0.03412616
0.03464323 0.03464323 0.03412616 0.0336091 0.0336091 0.0336091
0.0336091 0.0336091 0.03309204 0.03257497 0.03257497 0.03257497
0.03257497 0.03257497 0.03257497 0.03257497 0.03257497 0.03205791
0.03205791 0.03205791 0.03205791 0.03257497 0.03257497 0.03257497]
```

```
>>> Two-Layer DNN (Final) - ALL Test Misclassification Error:
[0.8858351 0.897463 0.86363636 0.89957717 0.79809725 0.76849894
0.71987315 0.63636364 0.602537 0.62156448 0.39112051 0.3435518
0.29598309 0.2653277 0.24735729 0.26638478 0.19556025 0.23361522
0.17019027 0.21247357 0.16384778 0.1744186 0.14693446 0.11839323
0.11733615 0.10676533 0.10887949 0.09830867 0.10042283 0.0961945
0.09196617 0.08773784 0.08668076 0.08668076 0.08668076 0.08562368
0.08773784 0.08668076 0.08350951 0.08245243 0.08139535 0.08139535
0.08139535 0.07928118 0.07928118 0.07928118 0.07928118 0.07928118
0.07716702 0.07610994 0.07505285 0.07293869 0.06976744 0.06976744
0.07082452 0.06976744 0.06976744 0.06976744 0.06976744 0.06871036
0.07082452 0.07188161 0.06976744 0.06976744 0.06976744 0.07082452
0.07082452 0.07188161 0.07082452 0.07082452 0.07082452 0.07082452
0.06976744 0.06976744 0.07082452 0.06871036 0.06765328 0.06659619
0.06659619 0.06448203 0.06448203 0.06342495 0.06342495 0.06448203
0.06553911 0.06553911 0.06236786 0.06131078 0.0602537 0.05919662
0.05919662 0.05919662 0.05813953 0.05813953 0.05813953 0.05813953
0.05813953 0.05708245 0.05708245 0.05708245 0.05708245 0.05602537
0.05602537 0.05602537 0.05602537 0.05602537 0.05602537 0.05602537
0.05708245 0.05708245 0.05708245 0.05708245 0.05708245 0.05708245
0.05708245 0.05602537 0.05602537 0.05602537 0.05602537 0.05496829
0.05391121 0.05391121 0.05391121 0.05391121 0.05391121 0.05496829
0.05285412 0.05285412 0.05285412 0.05285412 0.05285412 0.05179704
```

### *c) Evaluation of Trial Data Set:*

The two-layer DNN achieved a 0 % misclassification error rate on the trial set. This could be due to the model's balanced architecture, which provided sufficient capacity to learn the data effectively without overfitting. The small size of the trial set may have also contributed to the perfect performance.

### *CV Misclassification Errors:*

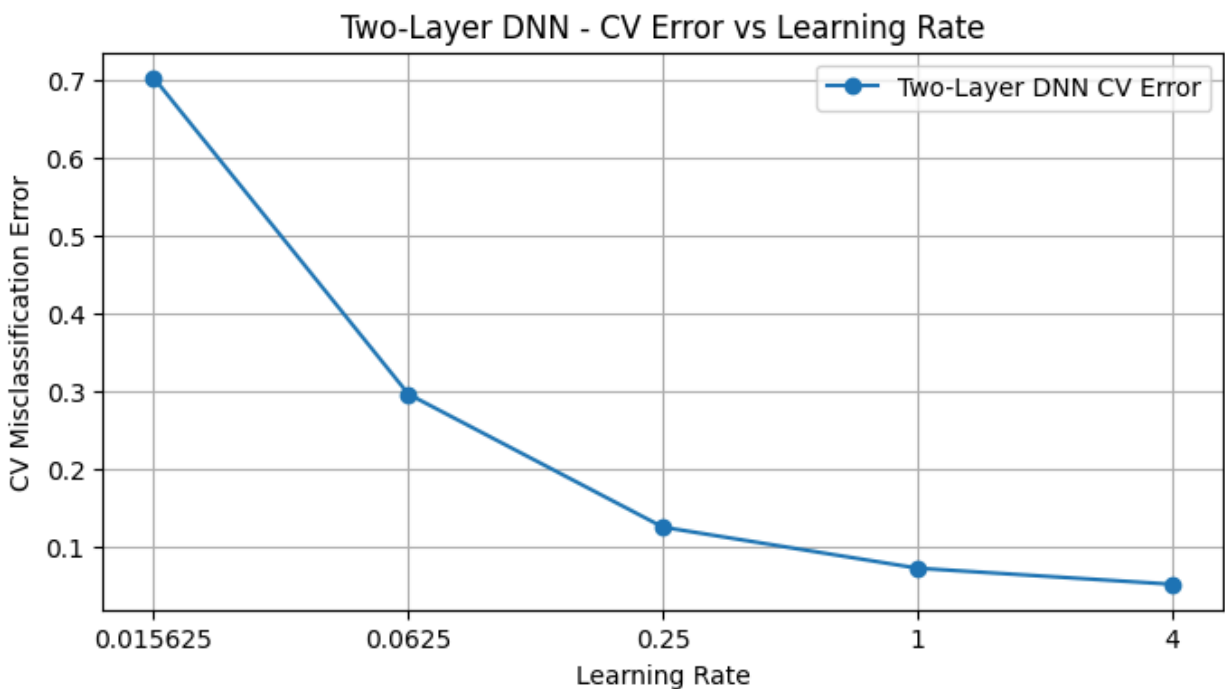
Based on 3-fold cross-validation, the best-performing learning rate for the Two-Layer DNN model was determined to be 4.

Learning Rate	Average CV Error
$4^{-3}$	0.7037
$4^{-2}$	0.2968
$4^{-1}$	0.1256

## CIS 481 Project 2: Handwritten Digits Recognition using Neural Networks

Written by: Zaynab Mourtada

$4^0$	0.0729
$4^1$	0.0522



### Learning Curves:

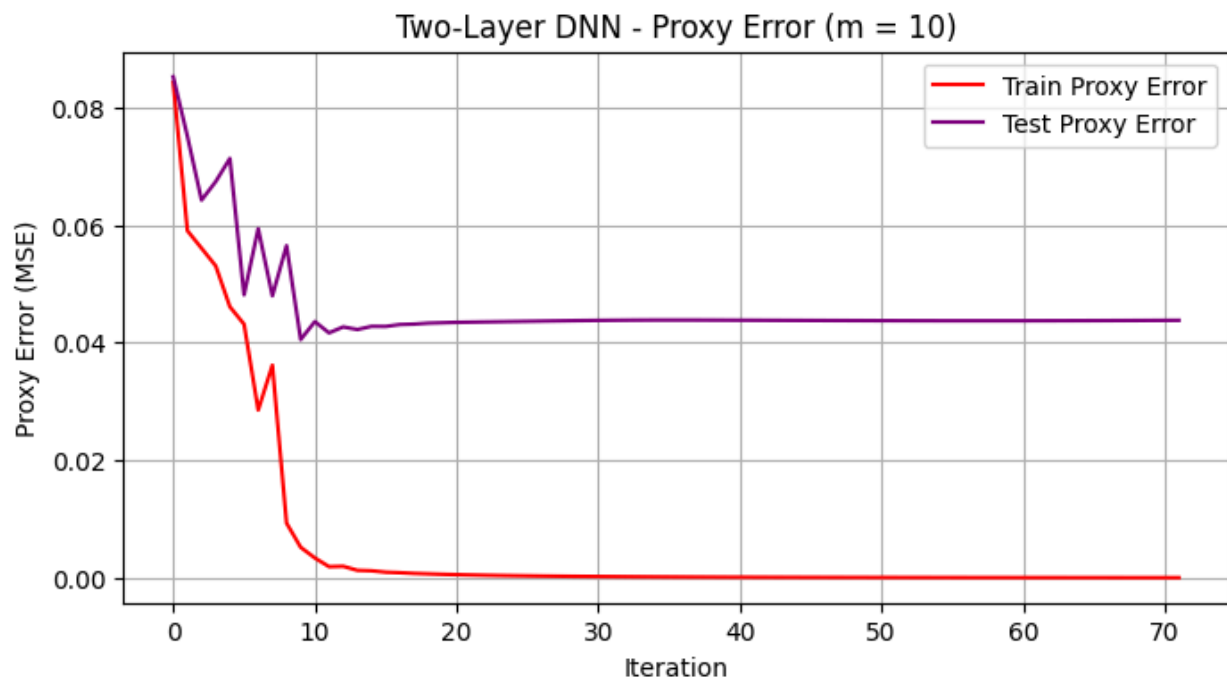
#### c) Learning Curves by Subset Size (Proxy & Misclassification Error vs Iterations)

This section shows how the two-layer DNN's proxy and misclassification errors changed over iterations for each training subset size.

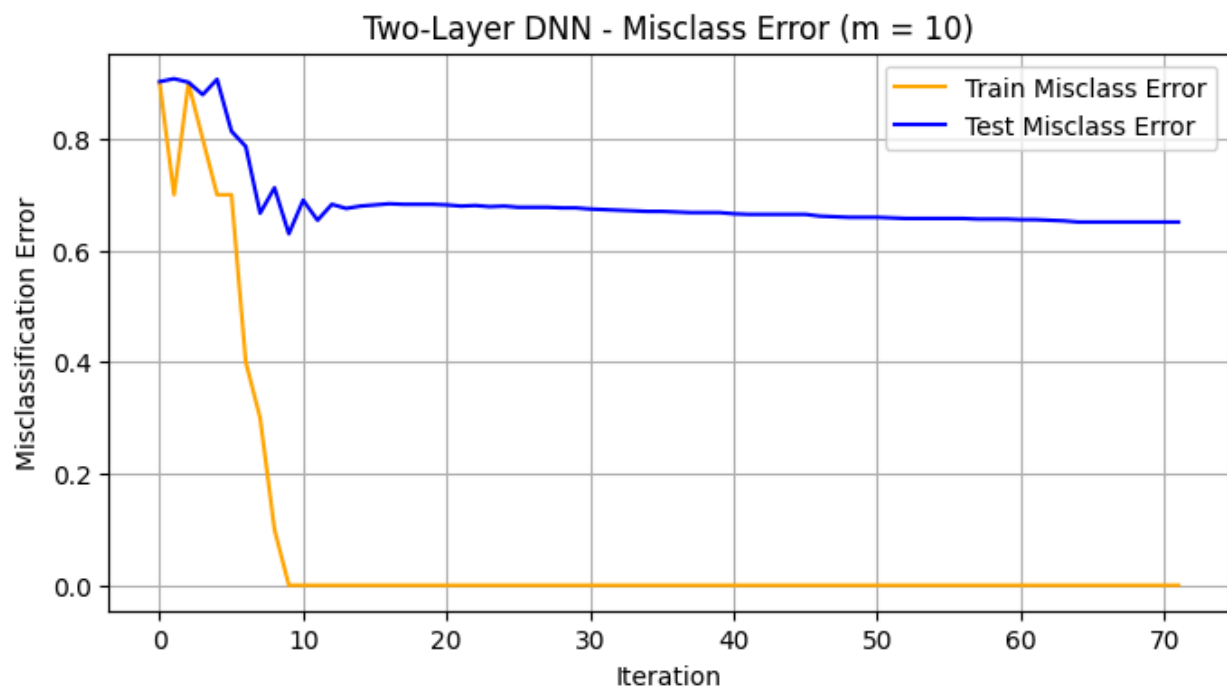
- $m = 10$

**CIS 481 Project 2: Handwritten Digits Recognition using Neural Networks**

**Written by: Zaynab Mourtada**



○

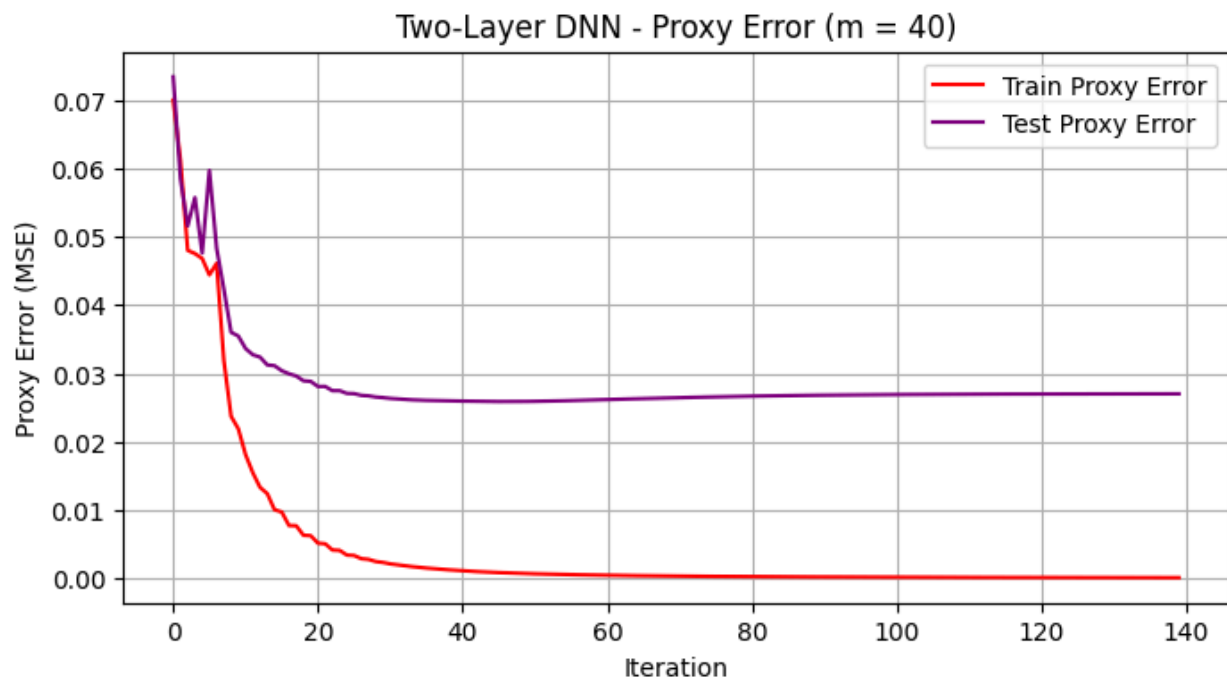


○

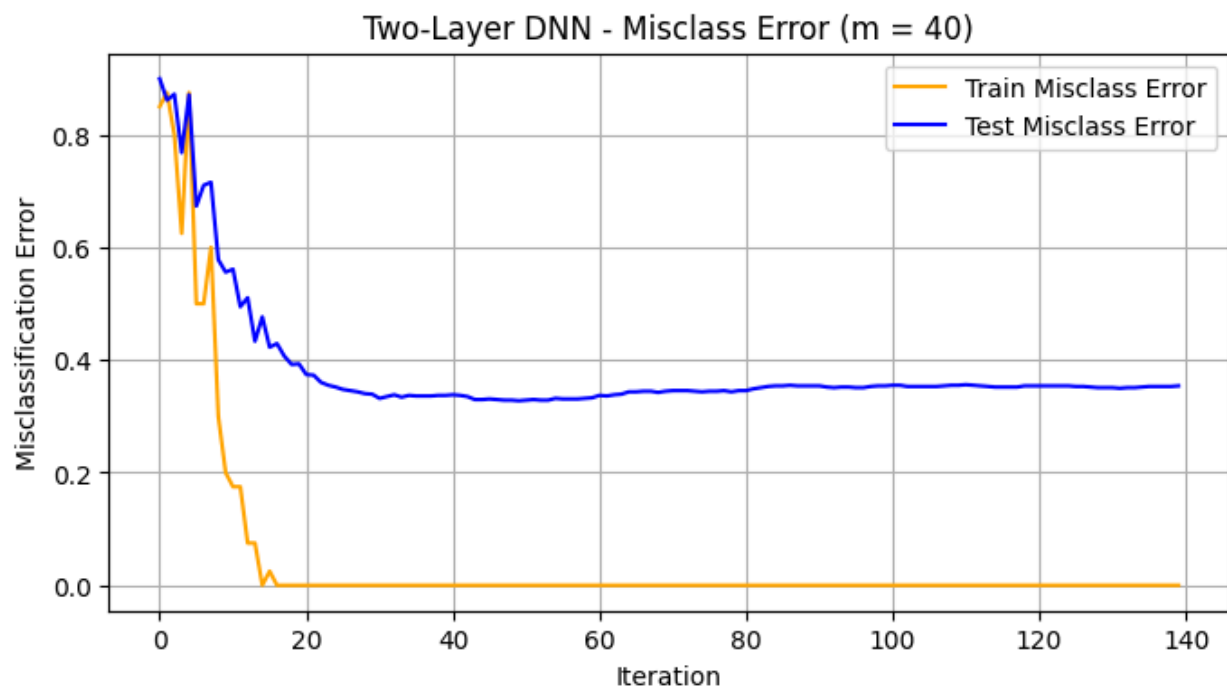
- m = 40

**CIS 481 Project 2: Handwritten Digits Recognition using Neural Networks**

**Written by: Zaynab Mourtada**



○

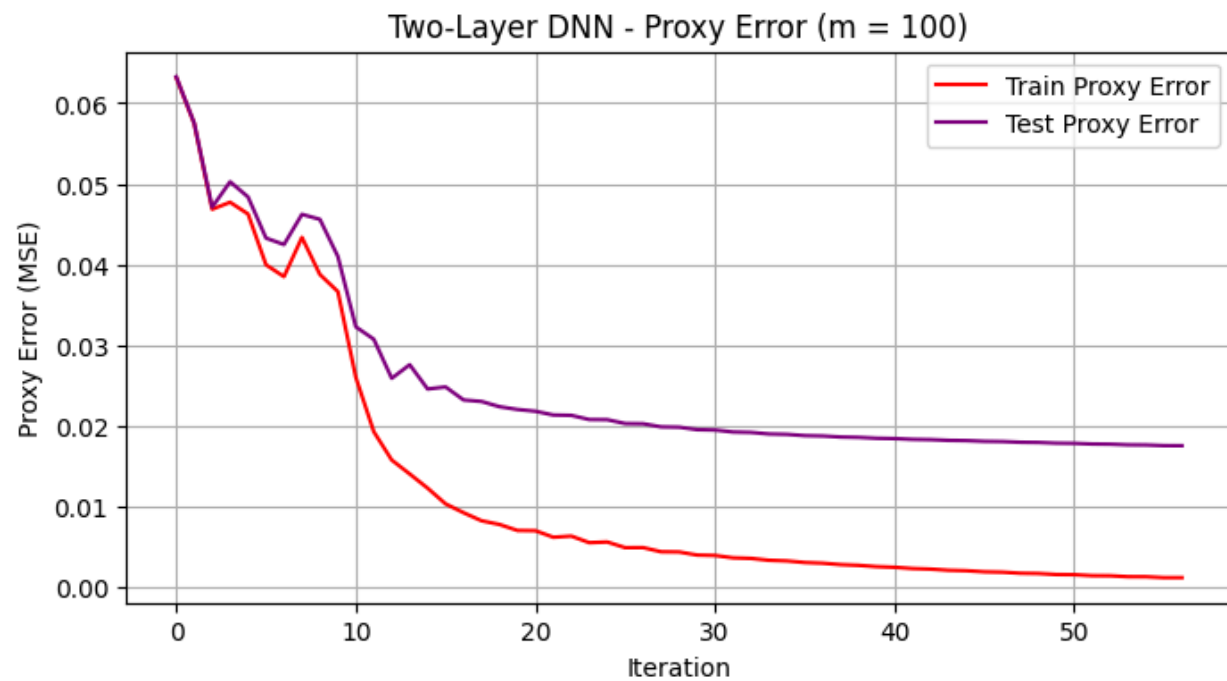


○

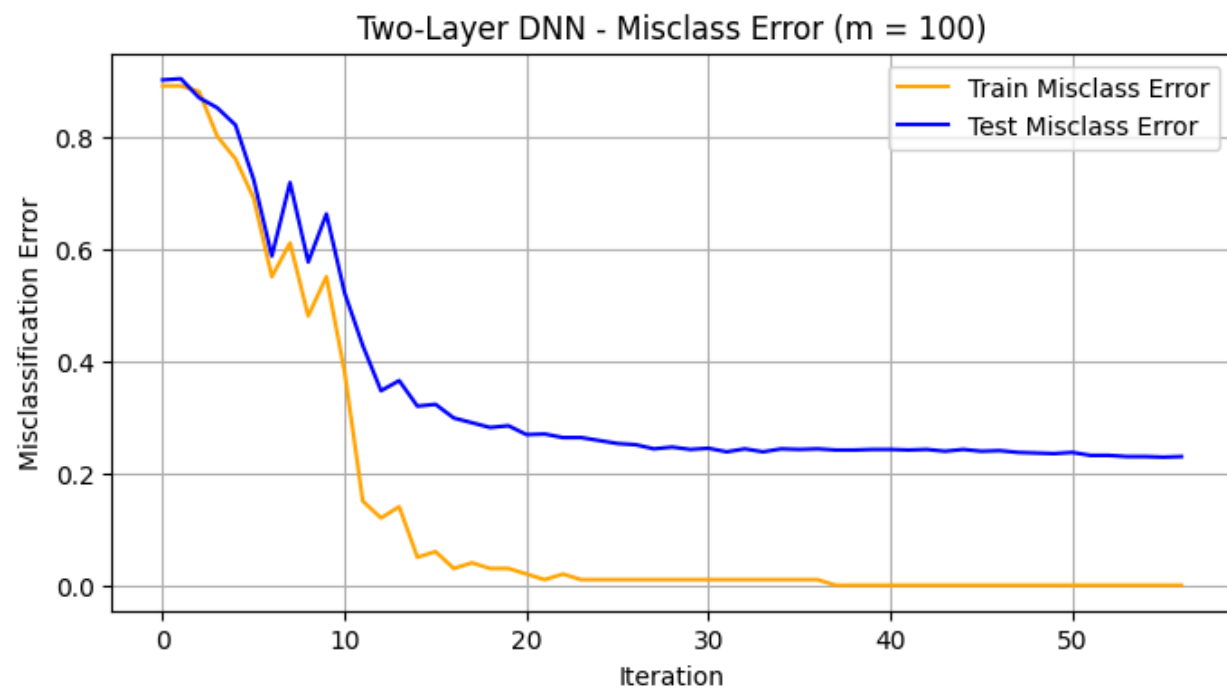
- m = 100

## CIS 481 Project 2: Handwritten Digits Recognition using Neural Networks

Written by: Zaynab Mourtada



○

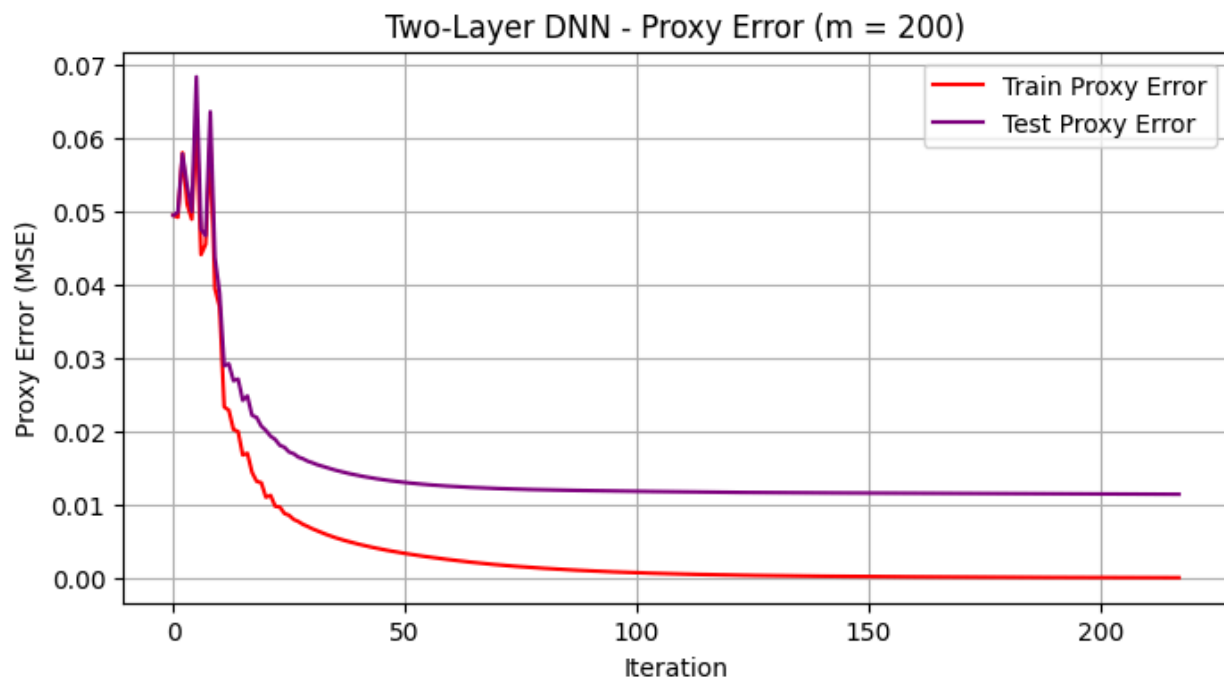


○

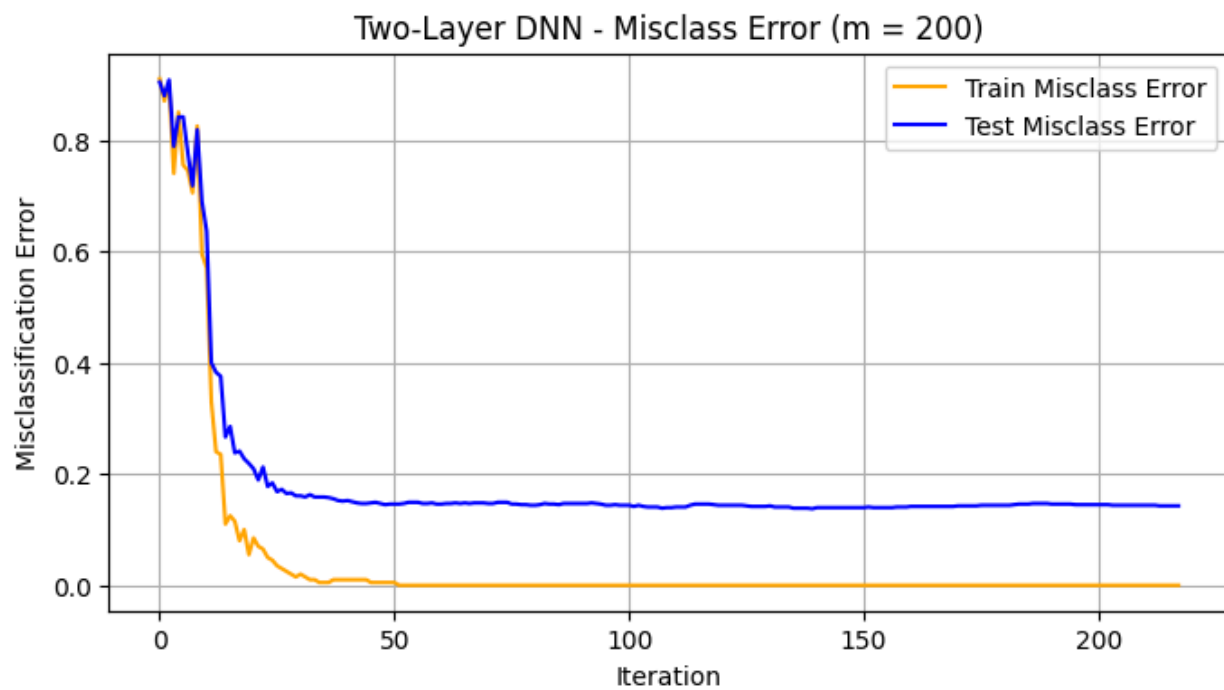
- m = 200

## CIS 481 Project 2: Handwritten Digits Recognition using Neural Networks

Written by: Zaynab Mourtada



○



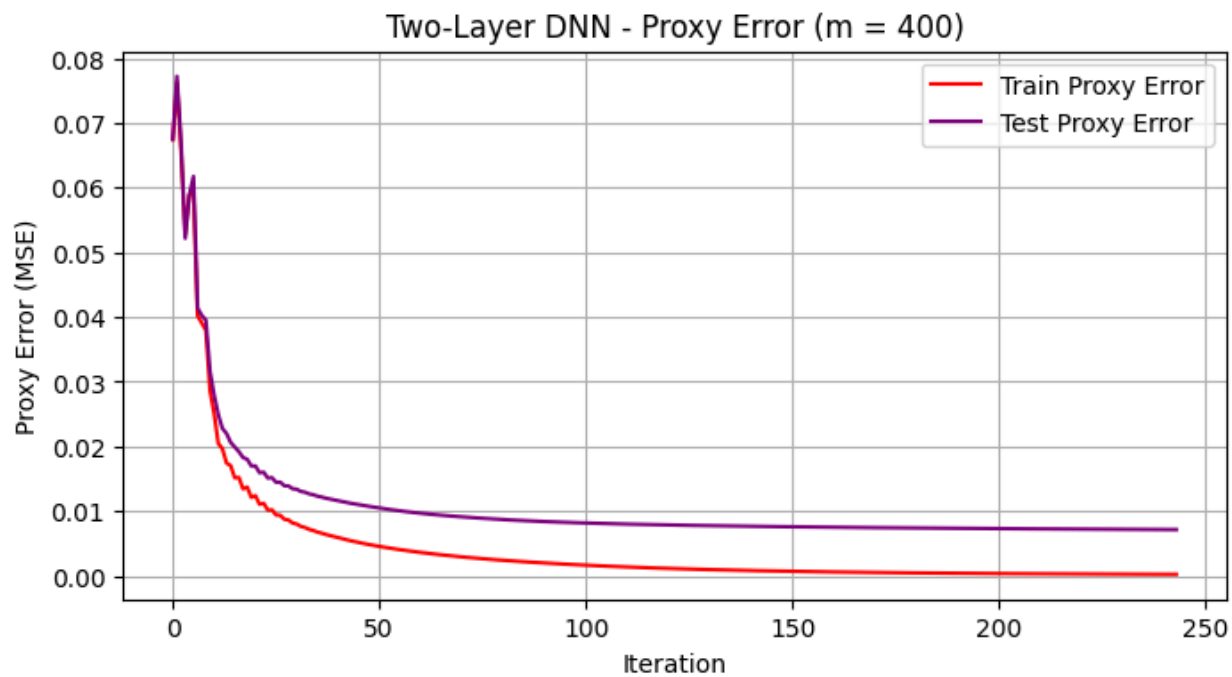
○

- m = 400

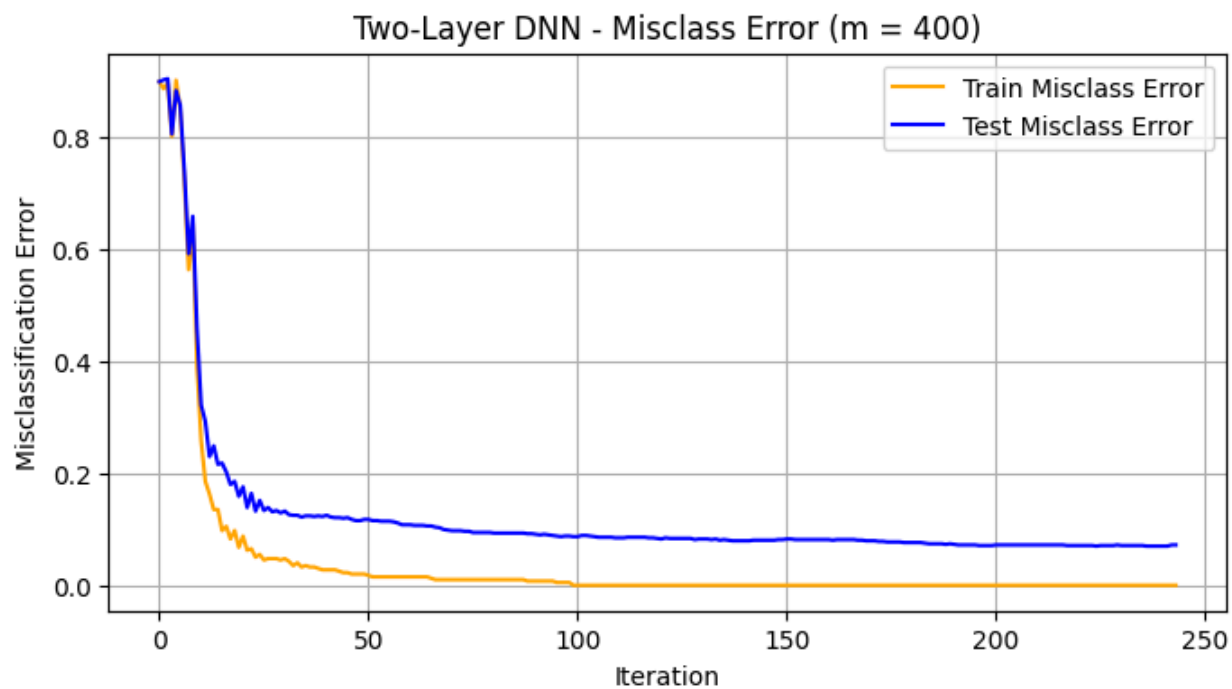


**CIS 481 Project 2: Handwritten Digits Recognition using Neural Networks**

**Written by: Zaynab Mourtada**



○

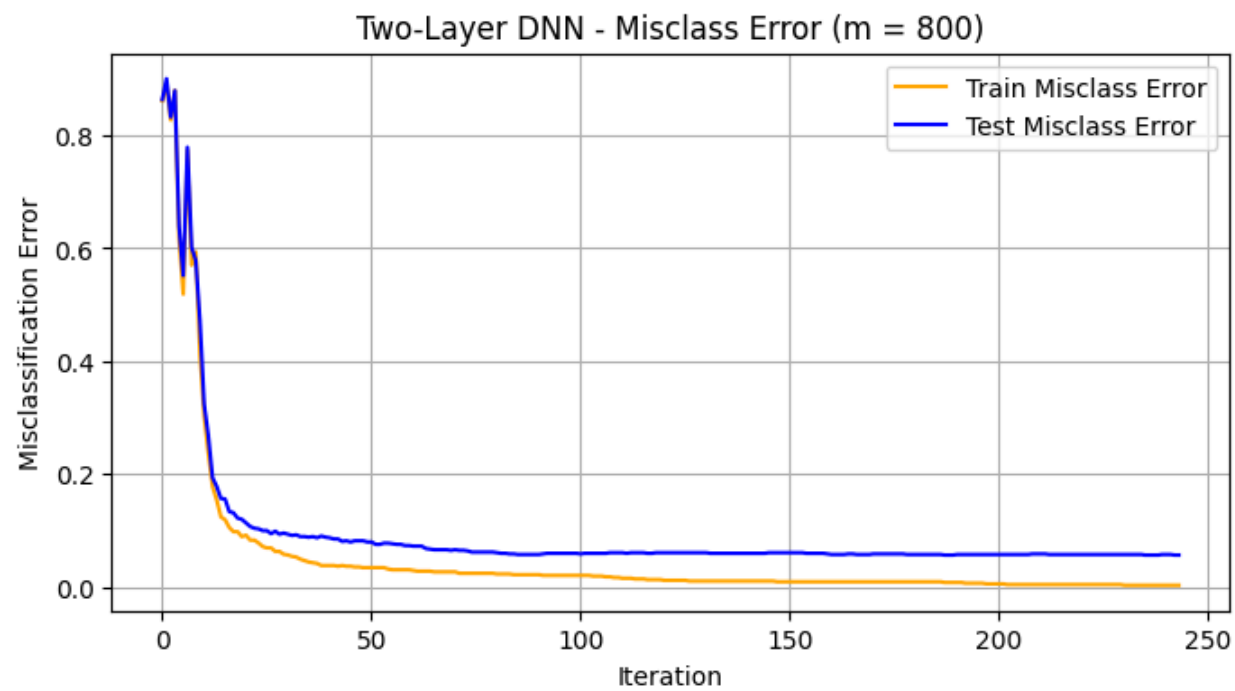
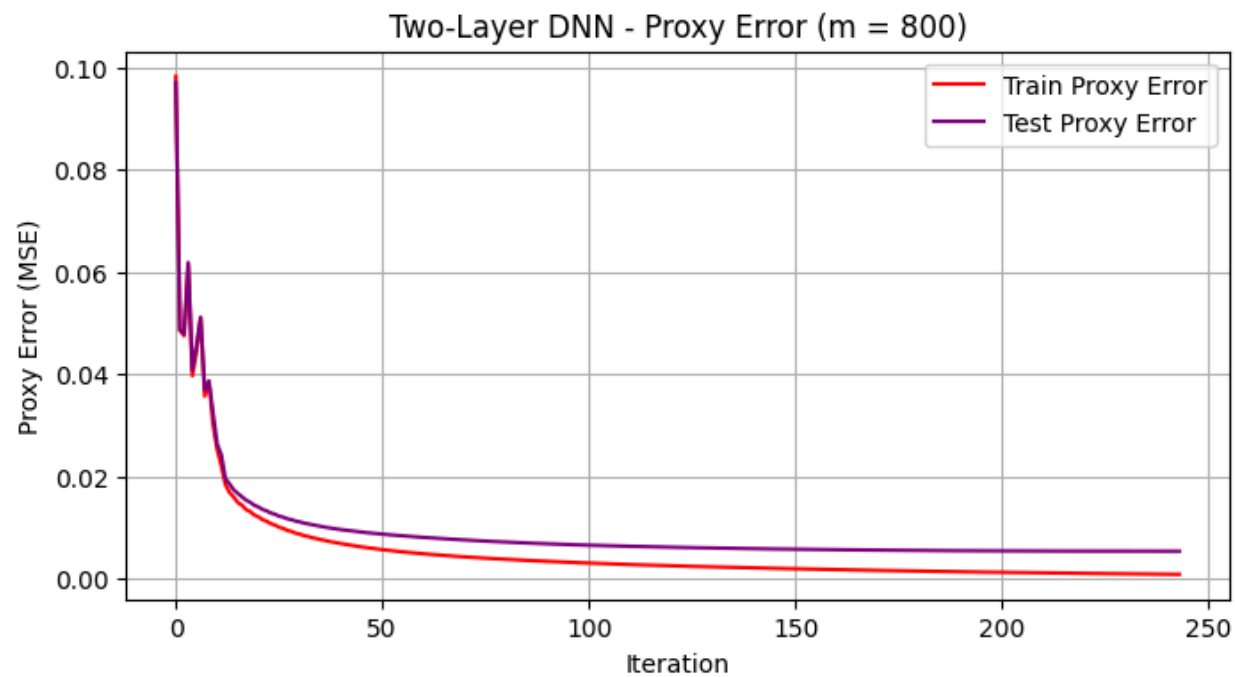


○

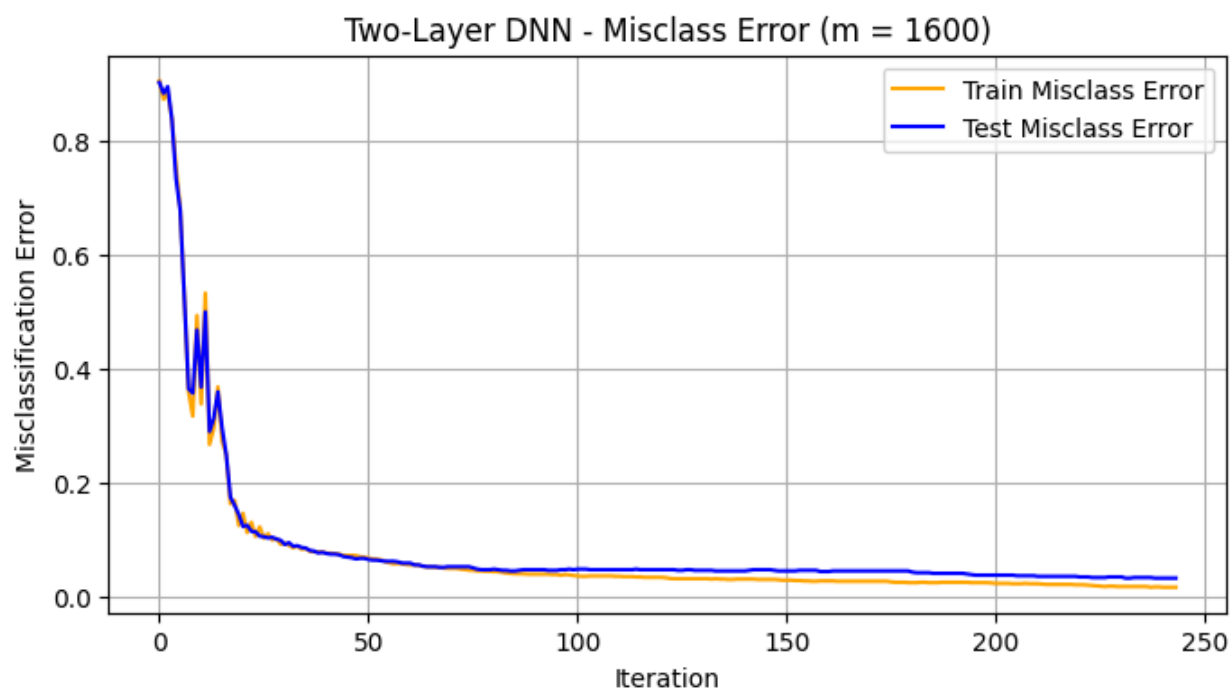
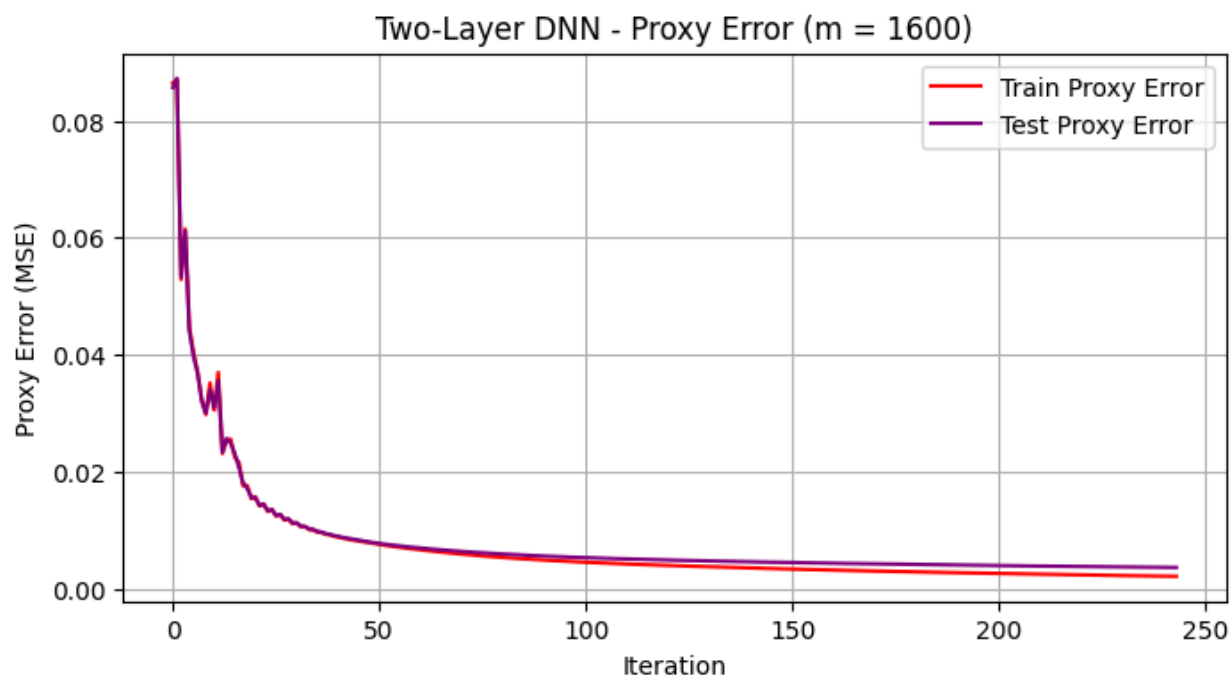
- m = 800

**CIS 481 Project 2: Handwritten Digits Recognition using Neural Networks**

**Written by: Zaynab Mourtada**

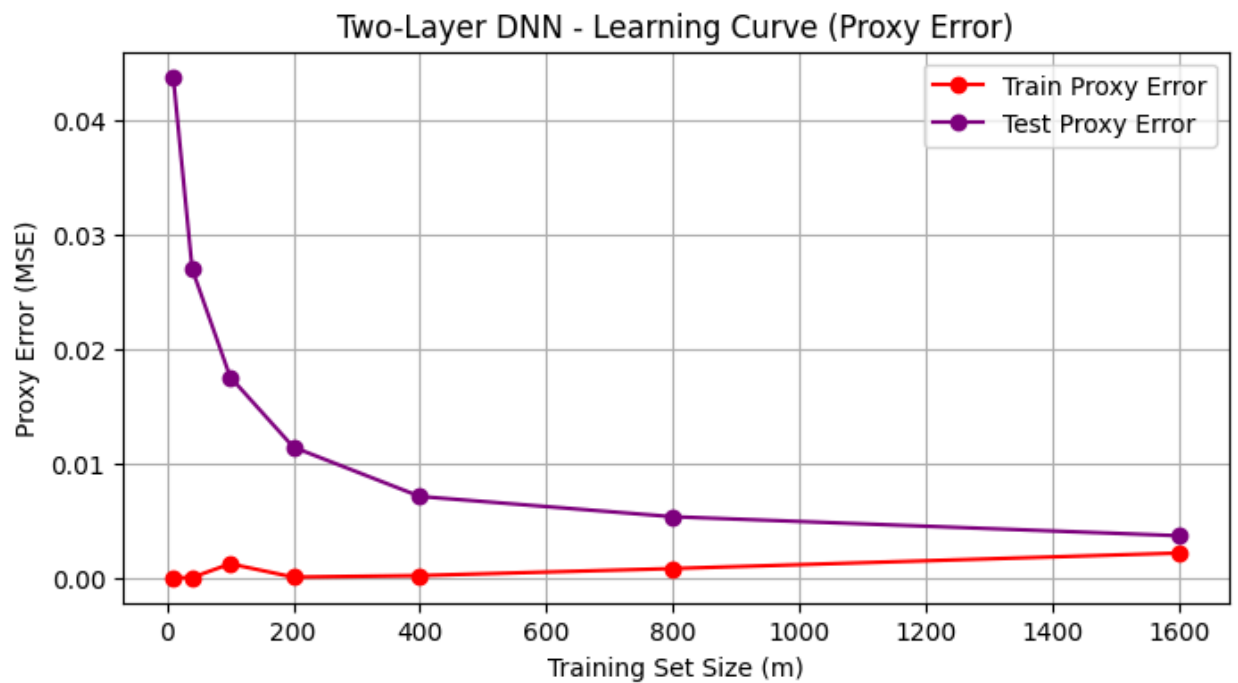


- m = 1600

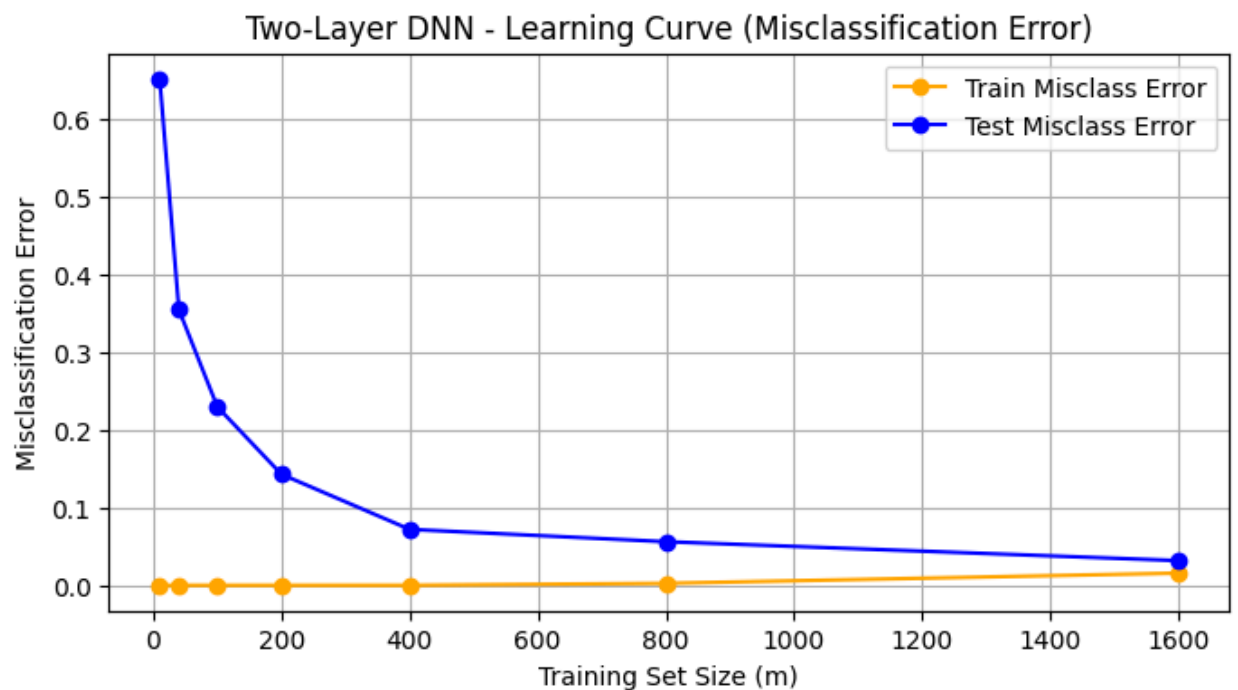


**d) Overall Learning Curves Across Subset Sizes**

As the training set size increases, the proxy error for both training and test sets steadily decreases, indicating that the two-layer DNN model continues to generalize better with more data. The test error drops sharply at first, then flattens out, showing diminishing returns but improved consistency.



As the training set size grows, the misclassification error drops sharply at first and then gradually levels off, especially on the test set. The training error stays very low throughout, while the gap between train and test narrows; this shows that the two-layer DNN generalizes better with more data and is not overfitting.



## ***Weight Parameter Interpretation***

This figure displays a sample of 10 randomly selected hidden unit weights from the first hidden layer of the two-layer DNN. Each visualization reflects the learned feature sensitivities of a specific unit. While some units show more defined patterns resembling digits or edges, others appear noisier, suggesting a range of specialization across the network.

