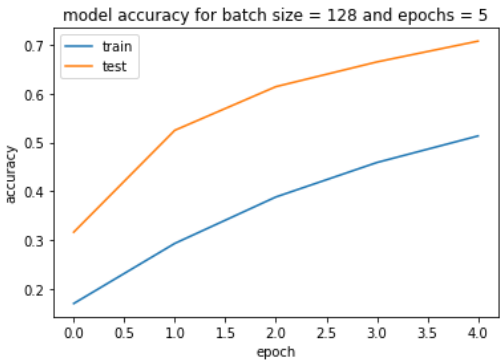
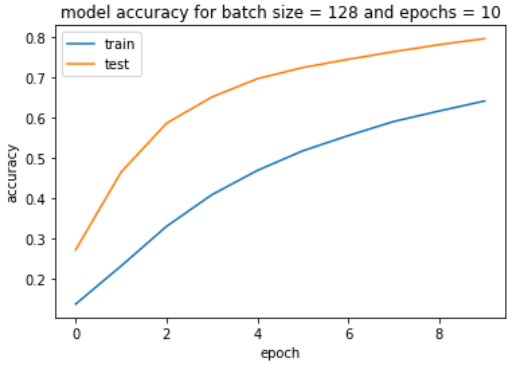
Tomas Kersulis
ENGRD 2720
December 7, 2021

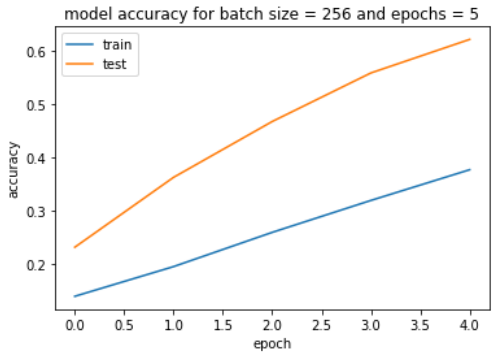
Deep Learning Assignment

Part A:
See attached code.

Part B:







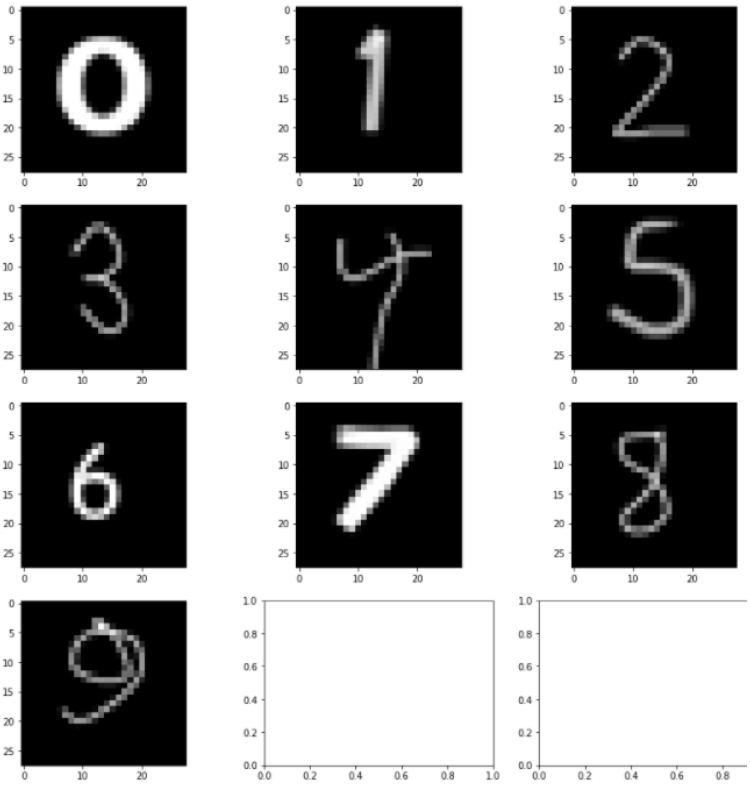
As the number of epochs goes up, the testing and training accuracy goes up. This means that the testing and training error goes down. However, there is a diminishing return as the number of epochs gets higher and higher. Although the accuracy is still increasing, the rate that the accuracy increases for each epoch as there are more epochs gets smaller and smaller. On the other hand, when the batch size increases, the testing and training accuracy goes down, as seen in the plot on the right: when the batch size doubled, the accuracy went down from the previous two plots. See attached code for implementation of plots and training.

Part C:

```
PREDICTION FOR 0:
[[0.31505567 0.00173791 0.13721776 0.01833169 0.12897553 0.04104533
  0.15454909 0.07527151 0.03356282 0.09425268]]
0

PREDICTION FOR 1:
[[0.03351057 0.2517618  0.10925283 0.09263696 0.04468814 0.09196073
  0.14955385 0.04684387 0.11990529 0.05988601]]
1

PREDICTION FOR 2:
[[0.07811872 0.12448416 0.15184996 0.14809339 0.04896225 0.10917827
  0.09565721 0.06445763 0.12762001 0.05157831]]
2

PREDICTION FOR 3:
[[0.05439913 0.1395092  0.11661524 0.15393469 0.06554987 0.10312691
  0.1113038  0.07685079 0.11379122 0.06491916]]
3

PREDICTION FOR 4:
[[0.05547984 0.10698614 0.05869912 0.08394239 0.09460792 0.1002872
  0.03540297 0.24279945 0.09117585 0.13061921]]
7

PREDICTION FOR 5:
[[0.17009987 0.02351675 0.13766229 0.14910895 0.04044463 0.1082163
  0.17262782 0.07604123 0.0896931  0.03258907]]
6

PREDICTION FOR 6:
[[0.0384681  0.0923771  0.04723219 0.02592868 0.21454568 0.09429953
  0.2276308  0.03314213 0.08649485 0.13988101]]
6

PREDICTION FOR 7:
[[0.02217666 0.24567519 0.25005102 0.18194996 0.0206135  0.05891413
  0.05077527 0.01493436 0.13715778 0.01775212]]
2

PREDICTION FOR 8:
[[0.05461684 0.14143714 0.12296331 0.12080595 0.06160023 0.1166417
  0.11545037 0.06008681 0.14482588 0.06157175]]
8

PREDICTION FOR 9:
[[0.08915431 0.06567582 0.18253653 0.13620092 0.07682996 0.07983834
  0.07623705 0.09302087 0.12612988 0.07437626]]
2
```

Images of the handwritten numbers converted to be 28x28 pixels and grayscale:

When I used 10 epochs of training and a batch size of 128, the neural network did a decent job of recognizing these drawings. It recognized 6/10 drawings: 0, 1, 2, 3, 6, and 8. The probabilities in the vector for each prediction are still pretty close together, indicating that it could use a little more training to increase the accuracy.



```
PREDICTION FOR 0:
[[0.11250687 0.08286638 0.10551096 0.0965279  0.09710749 0.09441981
  0.10300487 0.10453475 0.09563547 0.10788552]]
0

PREDICTION FOR 1:
[[0.08652361 0.11912126 0.10842708 0.09977197 0.08988219 0.09503496
  0.1017872  0.09758462 0.10253926 0.09932784]]
1

PREDICTION FOR 2:
[[0.09350544 0.10758257 0.11248199 0.10626546 0.09436122 0.09250274
  0.09755018 0.09228996 0.1063097  0.09715068]]
2

PREDICTION FOR 3:
[[0.09335554 0.10683256 0.10936598 0.10863285 0.09518518 0.09312108
  0.09705039 0.0981376  0.10363245 0.09468637]]
2

PREDICTION FOR 4:
[[0.09353901 0.10283856 0.10173243 0.10097447 0.09808213 0.09606276
  0.09238035 0.1082314  0.1042856  0.10187329]]
7

PREDICTION FOR 5:
[[0.11034217 0.09056734 0.10318546 0.10472263 0.09638595 0.092379
  0.10399913 0.09995797 0.09930774 0.09915257]]
0

PREDICTION FOR 6:
[[0.09470804 0.10267632 0.10395351 0.09061453 0.09604537 0.09344407
  0.10443516 0.09745742 0.10184505 0.11482058]]
9

PREDICTION FOR 7:
[[0.08649132 0.1207375  0.13337779 0.1156503  0.08809976 0.08435381
  0.08486852 0.08290834 0.11018447 0.0933281 ]]
2

PREDICTION FOR 8:
[[0.09410651 0.10921456 0.10975623 0.10661932 0.09219055 0.09035941
  0.09808027 0.09544943 0.10545179 0.09877199]]
2

PREDICTION FOR 9:
[[0.10042978 0.10310657 0.1112949  0.10273631 0.09251686 0.09223618
  0.0959544  0.09815247 0.10932368 0.09424886]]
2
```

When I redid the training with only 5 epochs and a batch size of 256, the neural network did a worse job than before, which was expected since the accuracy went down. This time the neural network was able to recognize 3/10 handwritten numbers: 0, 1, and 2. The predictions and their probabilities can be seen in the picture to the left.

When I increased the training to 20 epochs and a batch size of 64, the neural network recognized 5/10 handwritten digits, as seen in the picture to the right. This is surprising because with 10 epochs, the neural network recognized 6 of the handwritten digits. This may be because some of my digits were pretty illegible to the network and therefore did a worse job predicting them. However, the digits that were clear saw a much larger probability for that digit in the probability vector when there were 20 epochs. This means that the network was more confident in those predictions than when the network was trained with 10 epochs, as expected. Also, another possible explanation for this may be because of the diminishing returns of an extra epoch of training after 10 epochs. As seen in part b, as the number of epochs increases, each additional epoch adds less and less accuracy to the neural network. Therefore, there isn't as big of a change in accuracy of the network on the handwritten digits from 10 to 20 epochs as 5 to 10 epochs even though there is a bigger change in epochs from 10 to 20. Overall, the neural network pretty much acted as expected since it was pretty inaccurate in recognizing the handwritten digits with only 5 epochs, but once it increased to 10 and 20 it became relatively accurate. As the number of epochs increased, the network got more and more confident in the clear, legible digits, but struggled with the less legible digits.

```
PREDICTION FOR 0:
[[3.0013004e-01 1.3570515e-04 2.5274573e-02 9.8049815e-04 1.2106608e-01
  1.6606580e-03 9.7465053e-02 2.1850693e-01 2.3423063e-03 2.3243822e-01]]
0

PREDICTION FOR 1:
[[0.00667138 0.5881939  0.12081818 0.02343422 0.0043391  0.03090505
  0.0756248  0.01064521 0.10637259 0.03299564]]
1

PREDICTION FOR 2:
[[0.04831037 0.09368362 0.34044793 0.17480502 0.01192156 0.11750709
  0.05417651 0.01397483 0.12599778 0.01917523]]
2

PREDICTION FOR 3:
[[0.03964646 0.10070892 0.15280904 0.2604681  0.03550777 0.09673143
  0.1610826  0.02874477 0.08931755 0.03498335]]
3

PREDICTION FOR 4:
[[0.05901568 0.04483676 0.0174294  0.07012565 0.06388177 0.07039793
  0.01363571 0.45056638 0.06748904 0.1426217 ]]
7

PREDICTION FOR 5:
[[0.16830982 0.01298749 0.14142527 0.21037687 0.01161252 0.06372015
  0.3153552  0.04047089 0.02918665 0.00655513]]
6

PREDICTION FOR 6:
[[0.00924401 0.02812607 0.01708256 0.00367181 0.3764529  0.08999517
  0.2714713  0.01173631 0.07735153 0.11486843]]
4

PREDICTION FOR 7:
[[1.13427592e-03 3.09635084e-02 7.75782585e-01 6.13881983e-02
  4.30531625e-04 3.87053937e-03 1.41485385e-03 6.89145600e-05
  1.24204874e-01 7.41692027e-04]]
2

PREDICTION FOR 8:
[[0.02919308 0.11098949 0.19554827 0.11049728 0.03001937 0.13167982
  0.10431329 0.01642436 0.23621088 0.03512415]]
8

PREDICTION FOR 9:
[[0.04662102 0.044533   0.3512698  0.13824262 0.04407632 0.04151828
  0.04960722 0.03981908 0.18974034 0.0545723 ]]
2
```