Sarthak Chauhan

Professor Nada Attar

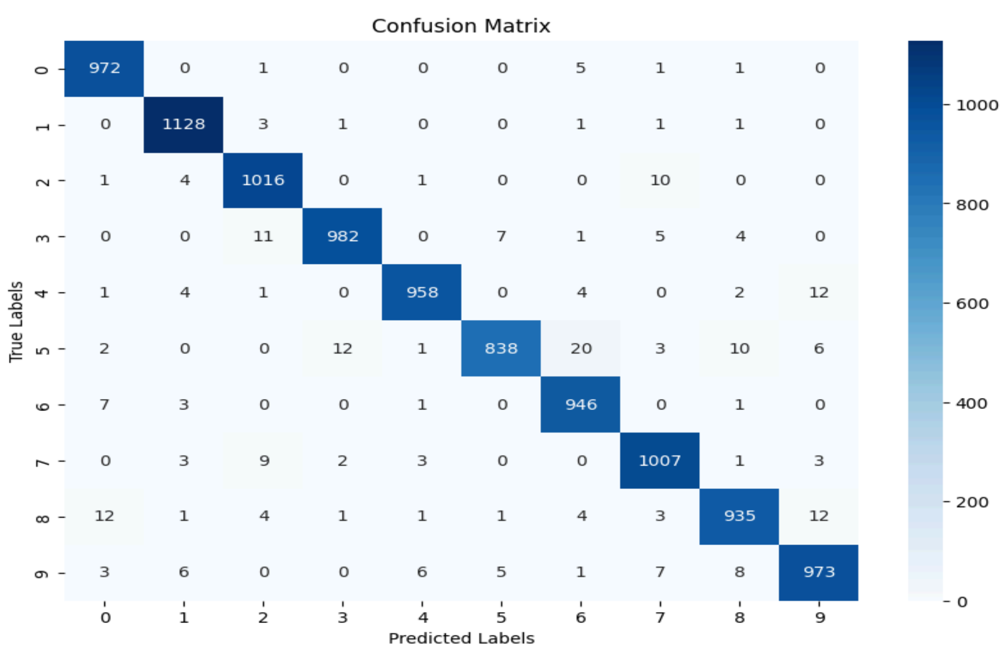
9th December 2024

CS 136 Optional Project

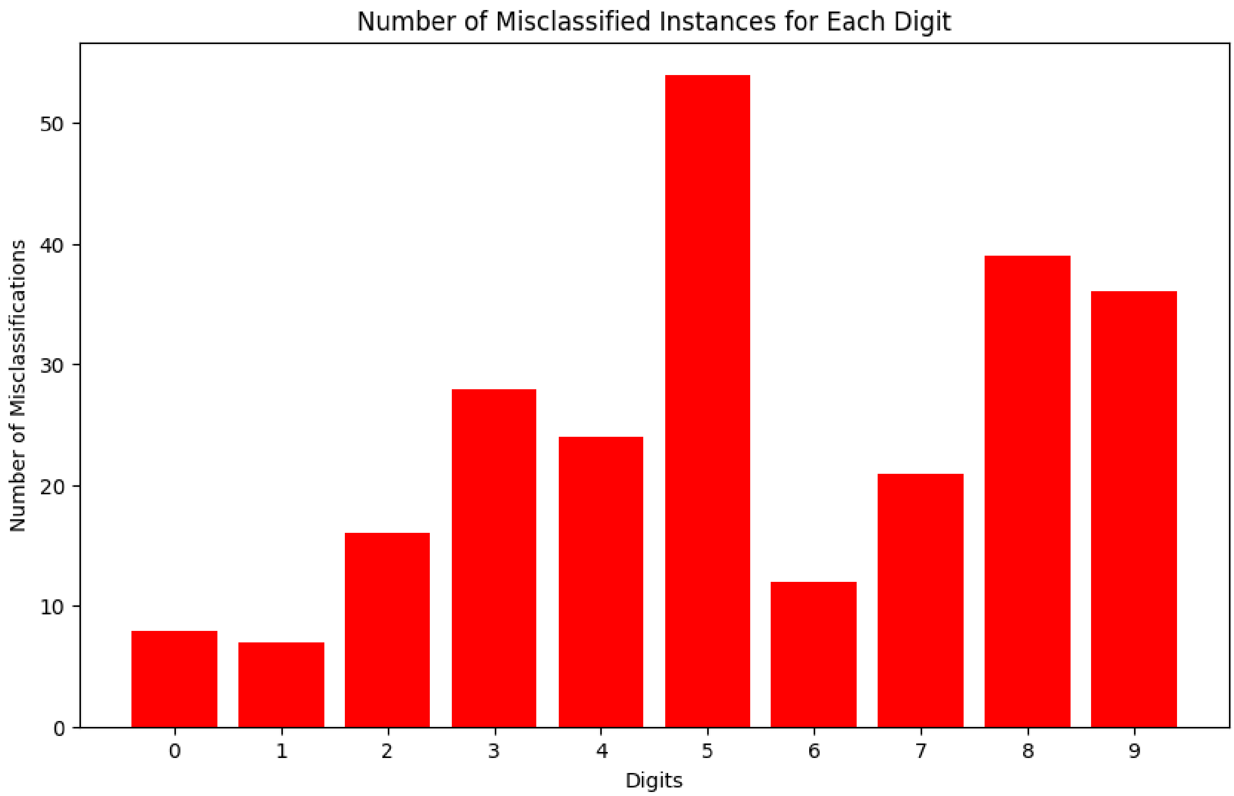
Exploring Dataset Bias in Computer Vision

1) **How consistent is the model trained on the base dataset? Is the base model better at some digits over the others? Why would this be?**

Overall the model was very consistent with the base dataset and showed good results when tested. We got a testing accuracy of `0.9754999876022339` which is very good and proves that the model trained is very consistent. I have also attached the confusion matrix to support my answer as you can see below.



Yes the base model is better at detecting some digits than others as you can see in the image below:



It detects the digit 0 and 1 with pretty good accuracy but struggles with numbers like 5, 8, 6 because of the way they are drawn and look pretty similar. For example the number “5” is similar to almost four other digits, hence our model cannot predict it correctly.

2) **Show the observed effects of dataset bias. For example, reduce the occurrence of “8” in the input data by 50%, and train the model and perform calculations on the model performance to determine how big an impact this skewed data made.**

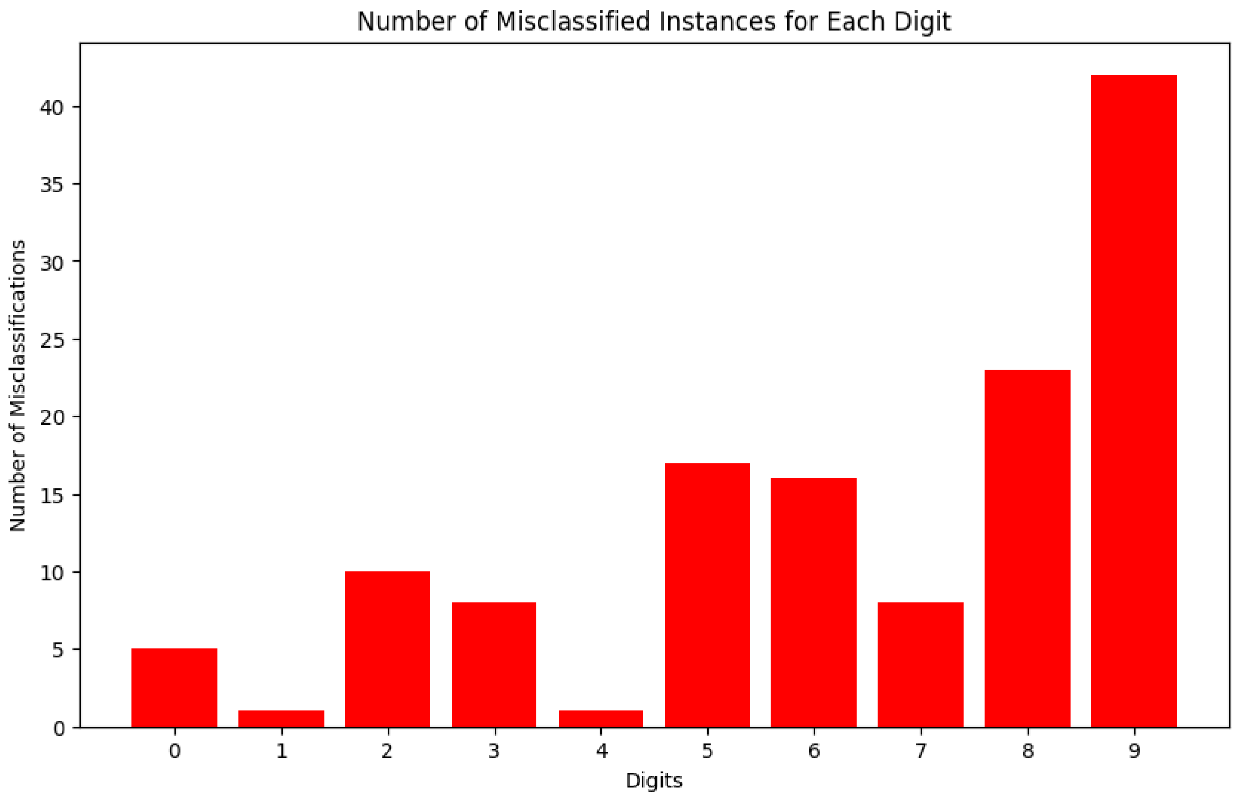
We will see the effect of reducing the occurrence of “8” by 50% and investigate the difference in accuracy, misclassification results.

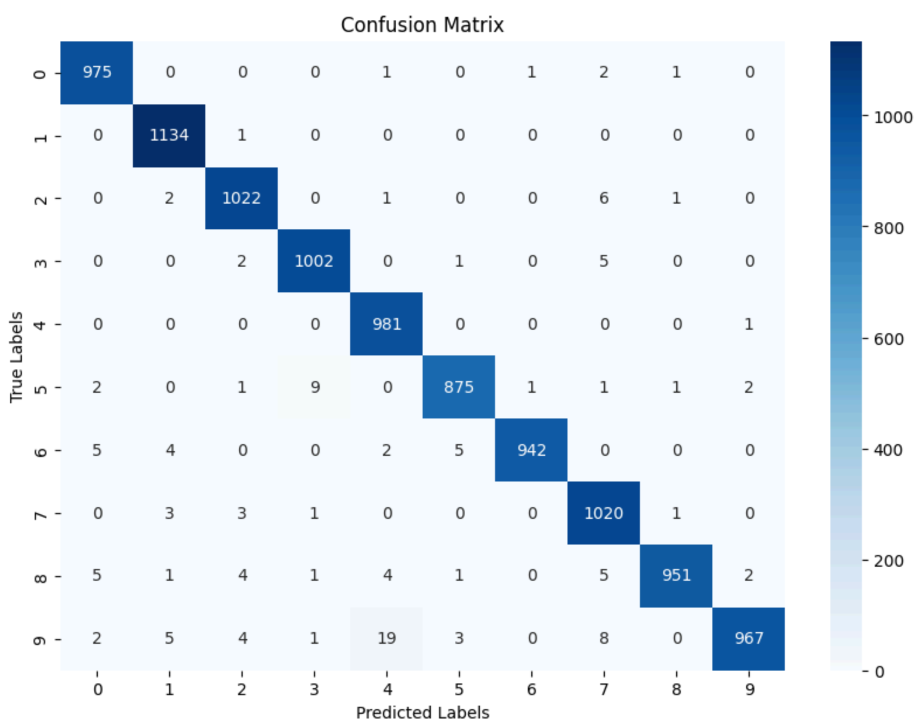
```
Original Dataset Distribution:
Digit 0: 5923
Digit 1: 6742
Digit 2: 5958
Digit 3: 6131
Digit 4: 5842
Digit 5: 5421
Digit 6: 5918
Digit 7: 6265
Digit 8: 5851
Digit 9: 5949

Biased Dataset Distribution:
Digit 0: 5923
Digit 1: 6742
Digit 2: 5958
Digit 3: 6131
Digit 4: 5842
Digit 5: 5421
Digit 6: 5918
Digit 7: 6265
Digit 8: 2984
Digit 9: 5949
```

As you can see above the number of occurrences in the biased dataset is less than 50% (2984 in biased dataset distribution). We saw that test loss even reduced further and the test accuracy has improved. **Test Loss: 0.04255188629031181, Test Accuracy: 0.9868999719619751**

But that is mostly due to the fact that the dataset has become more skewed than before, effectively simplifying the model's learning task.

Now if we see that number of misclassified images in the experiments **we will see that the digit “9” is misclassified the most. Because of the lower number of training dataset of digit “8”, our model never really gets better at learning digits similar to it. You can see the instances below as well.** I have attached the results for proof in the page below.





	0	1	2	3	4	5	6	7	8	9
0	975	0	0	0	1	0	1	2	1	0
1	0	1134	1	0	0	0	0	0	0	0
2	0	2	1022	0	1	0	0	6	1	0
3	0	0	2	1002	0	1	0	5	0	0
4	0	0	0	0	981	0	0	0	0	1
5	2	0	1	9	0	875	1	1	1	2
6	5	4	0	0	2	5	942	0	0	0
7	0	3	3	1	0	0	0	1020	1	0
8	5	1	4	1	4	1	0	5	951	2
9	2	5	4	1	19	3	0	8	0	967

3) **How much bias do you need to create to be able to see the effects in the data?**

The experiment reveals that the effects of dataset bias become significant when the reduction in the target digit's representation exceeds 30-50%. As the size of the training dataset is reduced the chances of our model failing to recognize the number decreases by a lot, although the model becomes more accurate the individual accuracy of recognizing that particular character reduces drastically. This demonstrates that even moderate levels of dataset bias can disproportionately disadvantage certain classes, emphasizing the critical need for balanced datasets in machine learning to ensure fair and robust model performance.

We need around 30-40% changes in the dataset to actually see the difference in our results.

4) **Does this change based on the size of the dataset?**

As mentioned above, yes this does change based on the size of the dataset. If we have a larger dataset, an increase/decrease in the frequency of a particular class will change our results drastically.

5) **Given a biased dataset, what do you think are some potential strategies to get a balanced, unbiased model?**

We will talk with respect to our current dataset of numbers and digits from the mnist dataset. There are couple of ways we can do it:

1) Data augmentation and transformations: We can apply transformations such as rotation, scaling, flipping, or adding noise to the existing samples. This artificially expands the dataset and helps the model generalize better.

2) Improving weights for our loss function: Assign higher weights to underrepresented classes in the loss function during model training. This penalizes the model more for misclassifying samples from minority classes, encouraging balanced performance.

6) **What are some ways biases could impact real-world applications of computer vision, and what are the potential implications.**

Real world example of biases impact can be:

Facial Recognition: Bias can cause unequal accuracy across demographics, leading to misidentifications or exclusions.

Medical Imaging: Underrepresentation of certain populations may result in misdiagnosis or overlooked conditions.

Autonomous Vehicles: Biased training data might make systems less effective in recognizing diverse environments or objects.

Retail and Surveillance: Biased models can unfairly target specific groups, raising ethical concerns and mistrust.

Hiring Systems: Computer vision tools for resume screening might favor certain groups, perpetuating inequalities.