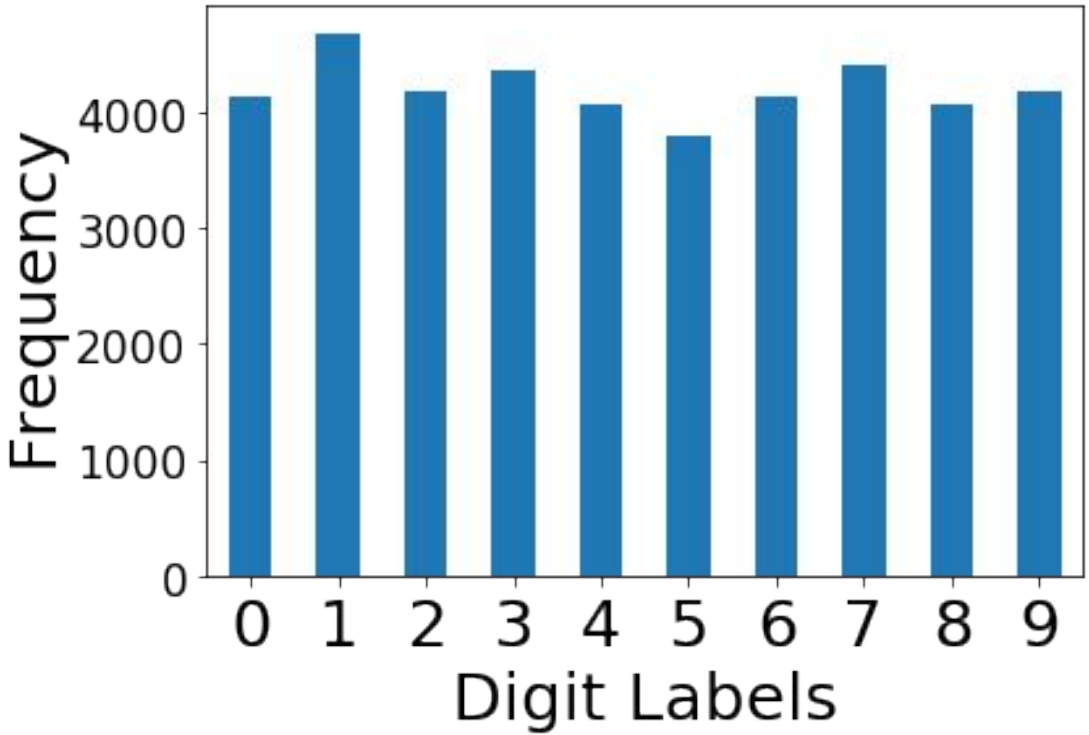
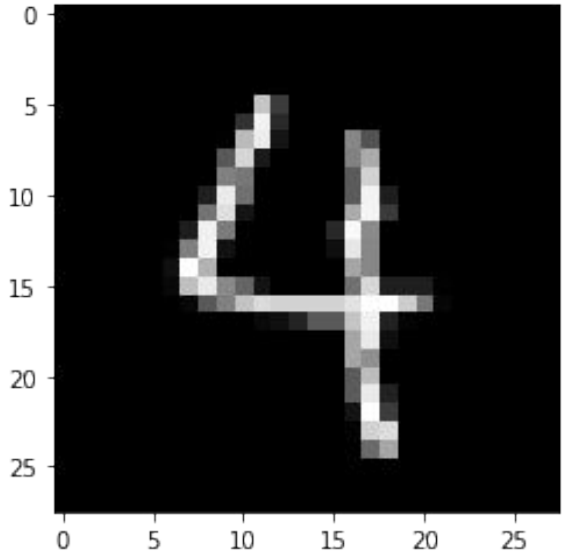
Use Case of Handwritten Digit Recognizer Machine Learning Models

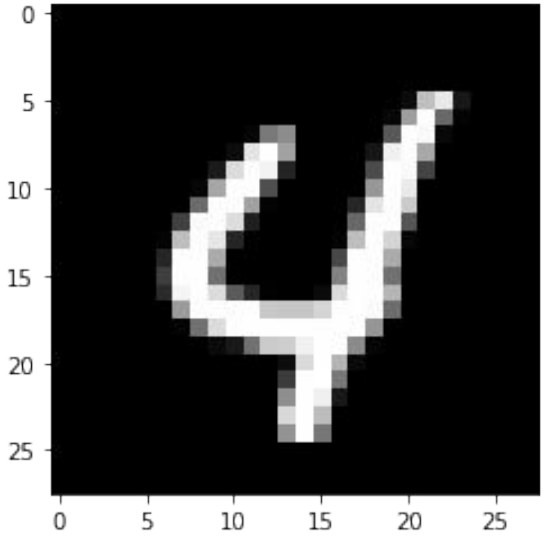
Ashish Piya

USPS Data Frequency



Digits in 28x28 matrix





Model prediction performance

- Accuracy of : 0.91
- Avg Precision: 0.91
- Avg Recall: 0.91

Use Cases of the Model

Add more features to model to classify alphabets

- Transcribing data
  - Healthcare information data
  - Book-keeping / account data

Transcribing Data

Companies can use machine learning model to digitize data

Eg: Cloudfactory

Cloudfactory and Spokeo

- 200 Million record (1790-1940)
- CF paid around $700k


Any
Questions ?

Appendix 1: Model Prediction Performance

Accuracy of : 0.91

Avg Precision: 0.91

Avg Recall: 0.91

Avg F-1 score: 0.91

Classification Report			
Digits	precision	recall	F1-score
0	0.95	0.96	0.96
1	0.97	0.95	0.96
2	0.89	0.91	0.9
3	0.9	0.87	0.89
4	0.95	0.88	0.91
5	0.96	0.81	0.88
6	0.96	0.95	0.96
7	0.74	0.93	0.82
8	0.82	0.92	0.87
accurac	0.91		

Appendix 2: Model Prediction Performance

Confusion Matrix										
row/col	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
0	761	0	7	5	0	1	5	0	10	0
1	0	917	18	3	4	0	3	1	20	3
2	3	3	734	18	4	0	4	1	43	0
3	6	2	18	792	2	8	0	4	68	12
4	1	1	16	2	744	1	9	7	17	51
5	13	1	7	33	5	606	7	0	65	13
6	8	1	1	0	1	12	786	0	15	0
7	3	2	9	2	10	0	0	783	7	82
8	4	13	10	17	2	1	4	0	723	6
9	4	1	4	4	13	0	0	23	14	756