



CS671 Deep Learning
Even Sem 2018-19
March 31th, 2019

Layer API

Assignment - 2

Submitted By

Roll No	Name
---------	------

B16066	Nikhil T R
B16021	Kaustubh Verma
B16004	Aj R Laddha

Contents

1	MNIST Dataset	1
1.1	Variation 1	1
1.2	Variation 2	2
1.3	Variation 3	3
2	Line Dataset	5
2.1	Variation 1	5
2.2	Variation 2	6
2.3	Variation 3	6
3	Inference	8

Chapter 1

MNIST Dataset

1.1 Variation 1

Details

Architecture	300-100-10	Testing Accuracy	0.9748	S. Deviation	0.01
Epochs	35	Batch Size	200	Learning Rate (SGD)	0.3

Table 1.1: Details

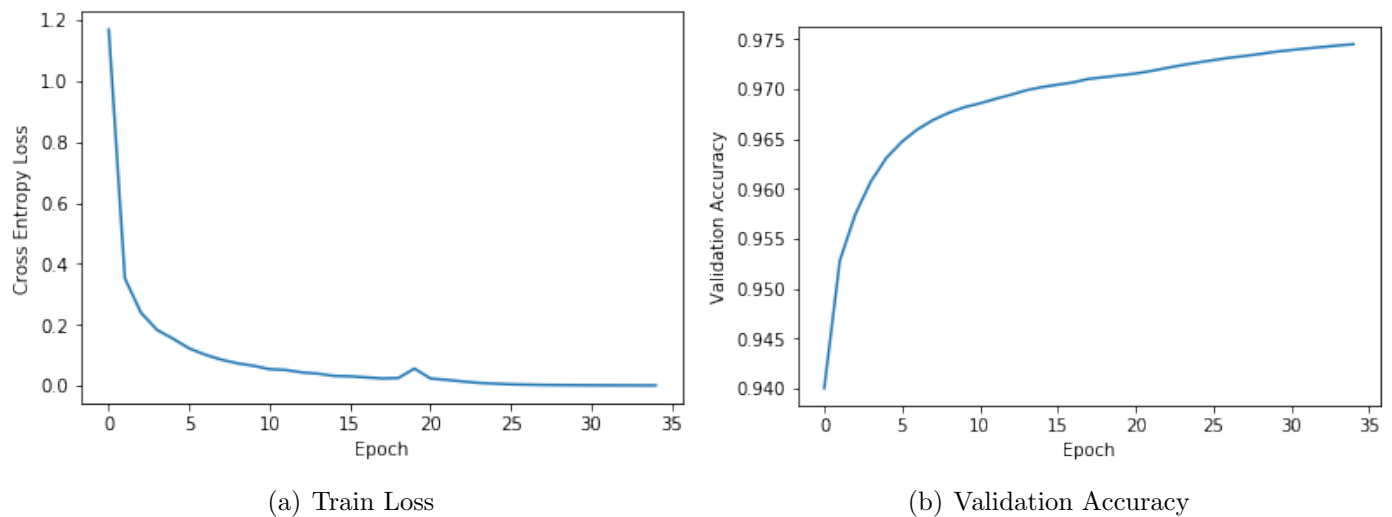


Figure 1.1: Training

968	0	1	0	0	3	3	2	2	1
1	1126	2	1	0	2	0	0	3	0
2	1	1008	3	3	0	2	9	4	0
2	0	2	993	0	2	0	3	5	3
0	0	3	1	957	1	4	2	1	13
2	0	0	8	1	870	4	1	4	2
4	3	1	1	1	7	940	0	1	0
0	2	5	3	0	0	0	1006	4	8
2	0	2	5	3	3	3	3	949	4
3	2	0	4	8	1	0	3	2	986

Table 1.2: Confusion Matrix

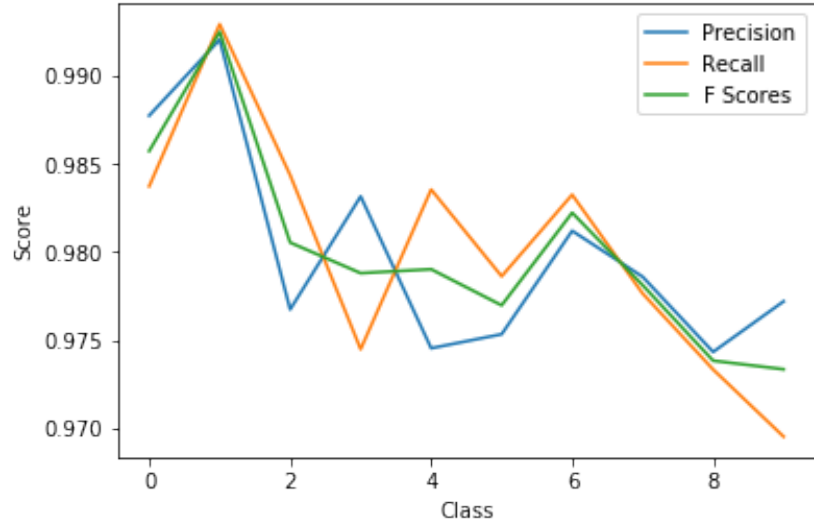


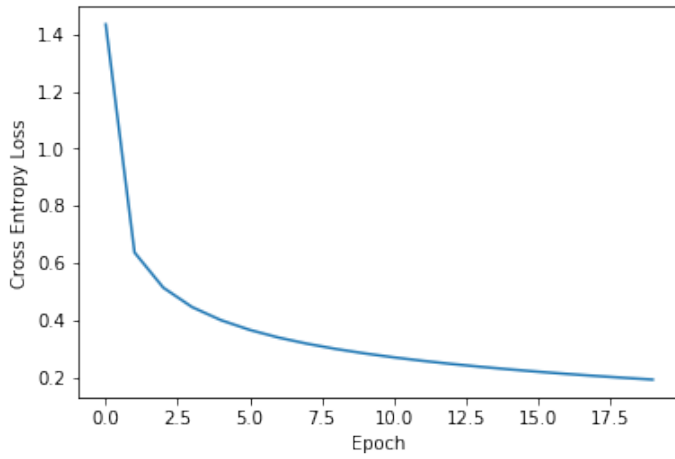
Figure 1.2: Scores

1.2 Variation 2

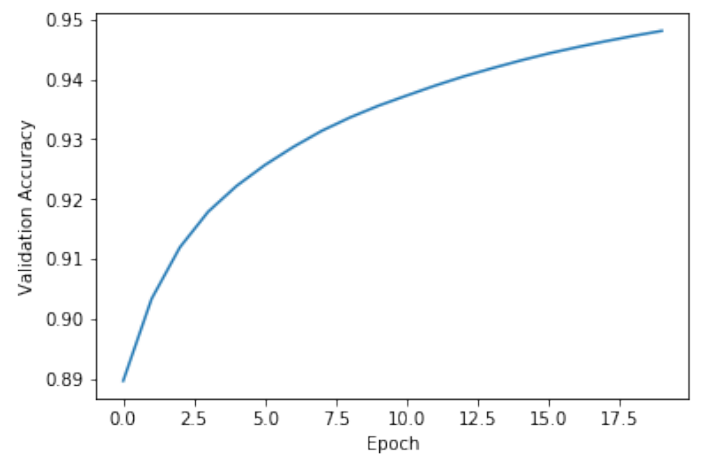
Details

Architecture	300-100-10	Testing Accuracy	0.9487	S. Deviation	0.08
Epochs	20	Batch Size	500	Learning Rate (SGD)	0.01

Table 1.3: Details



(a) Train Loss



(b) Validation Accuracy

Figure 1.3: Training

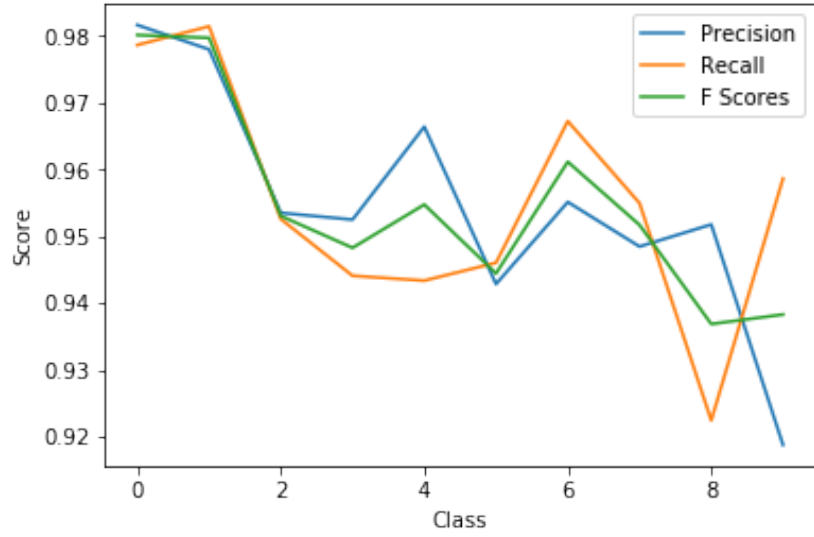


Figure 1.4: Scores

962	0	5	0	0	4	3	2	3	1
0	1110	3	3	0	1	3	1	13	1
4	2	984	12	6	0	3	11	9	1
0	0	10	962	0	12	0	9	11	6
1	0	7	0	949	0	5	4	5	11
5	0	1	19	3	841	11	1	8	3
6	3	6	1	10	10	915	0	7	0
0	8	12	8	6	3	0	975	5	11
2	2	5	7	5	10	5	5	927	6
3	6	0	7	27	8	1	13	17	927

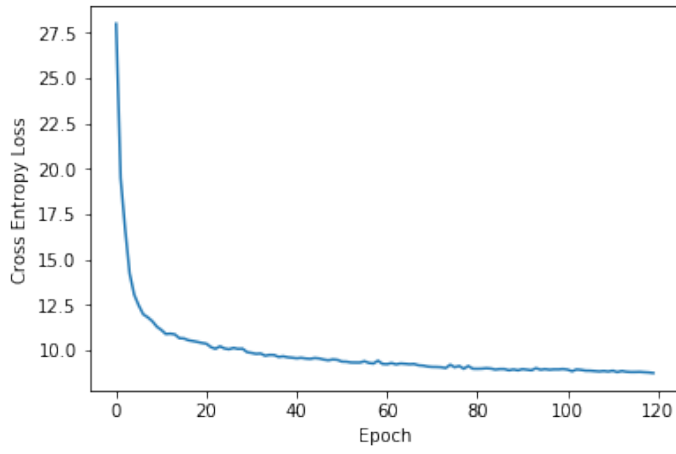
Table 1.4: Confusion Matrix

1.3 Variation 3

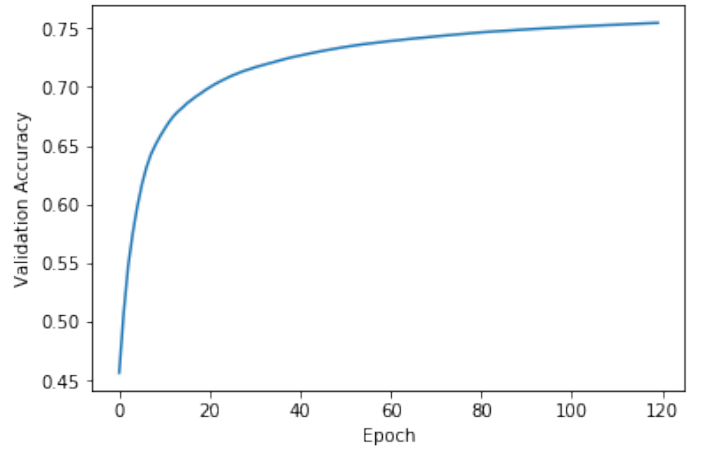
Details

Architecture	300-100-10	Testing Accuracy	0.9487	S. Deviation	1.0
Epochs	120	Batch Size	50	Learning Rate (SGD)	0.005

Table 1.5: Details



(a) Train Loss



(b) Validation Accuracy

Figure 1.5: Training

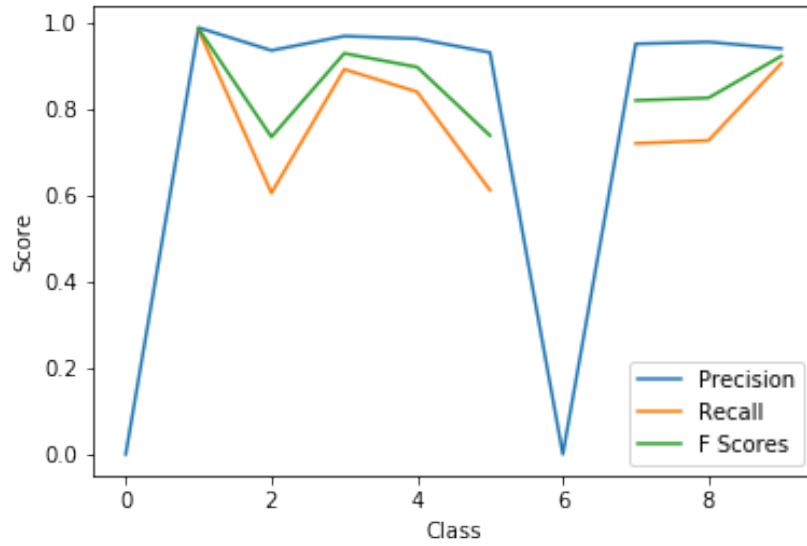


Figure 1.6: Scores

0	0	161	24	6	399	0	314	55	21
0	1120	5	5	0	2	0	1	2	0
0	0	964	22	4	1	0	20	18	3
0	0	8	977	2	5	0	5	9	4
0	1	8	0	944	0	0	2	5	22
0	1	4	26	1	829	0	5	16	10
0	7	430	9	145	111	0	8	230	18
0	6	10	7	6	2	0	976	6	15
0	0	4	17	3	4	0	10	929	7
0	6	1	10	15	4	0	16	10	947

Table 1.6: Confusion Matrix

Chapter 2

Line Dataset

2.1 Variation 1

Details

Architecture	192-96	Testing Accuracy	0.9741	S. Deviation	0.03
Epochs	30	Batch Size	25	Learning Rate (SGD)	0.3

Table 2.1: Details

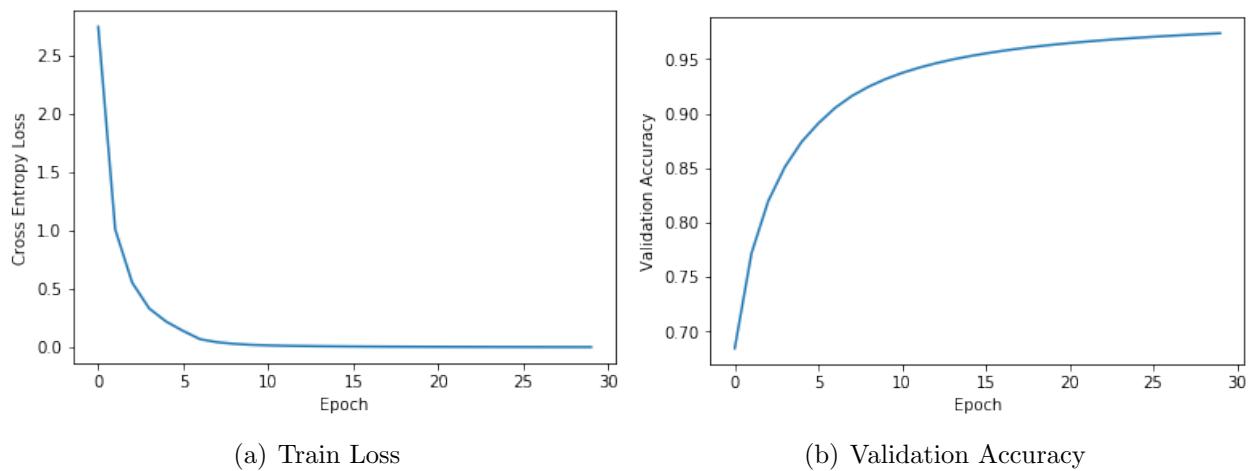


Figure 2.1: Training

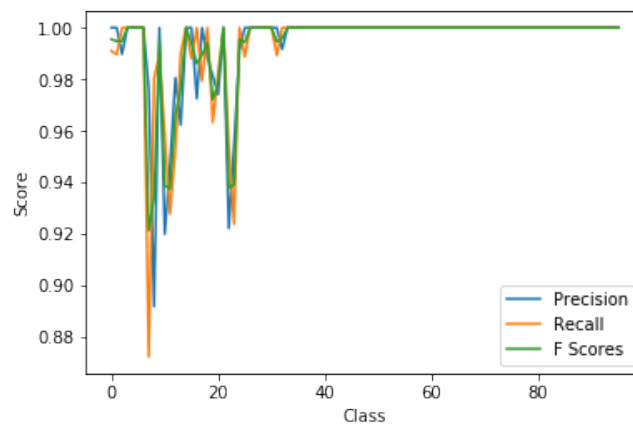


Figure 2.2: Scores

2.2 Variation 2

Details

Architecture	128-96	Testing Accuracy	0.7334	S. Deviation	0.1
Epochs	50	Batch Size	800	Learning Rate (SGD)	0.3

Table 2.2: Details

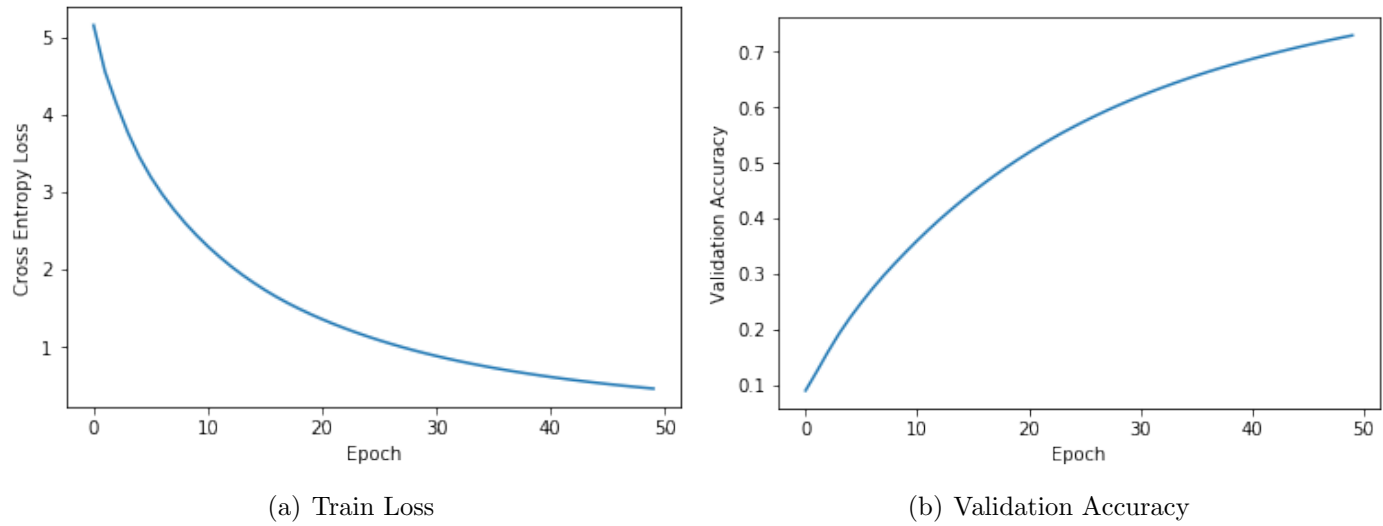


Figure 2.3: Training

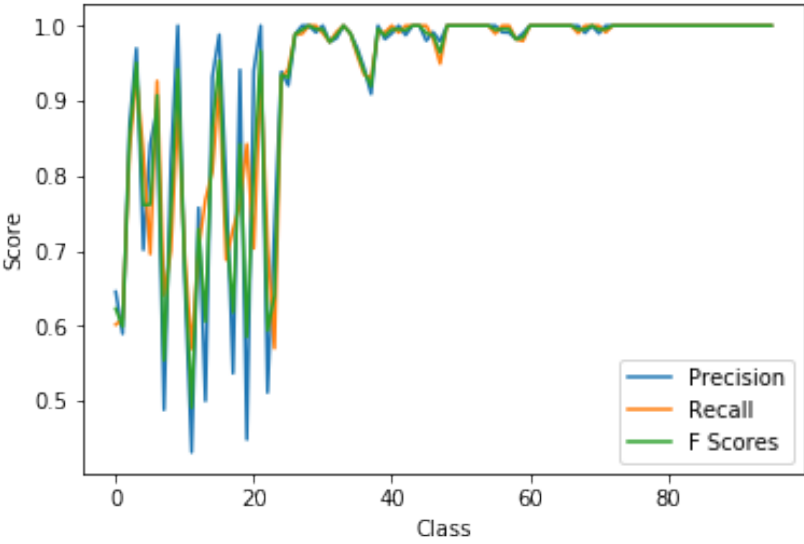


Figure 2.4: Scores

2.3 Variation 3

Details

Architecture	256-128-96	Testing Accuracy	0.6536	S. Deviation	1.0
Epochs	50	Batch Size	50	Learning Rate (SGD)	0.1

Table 2.3: Details

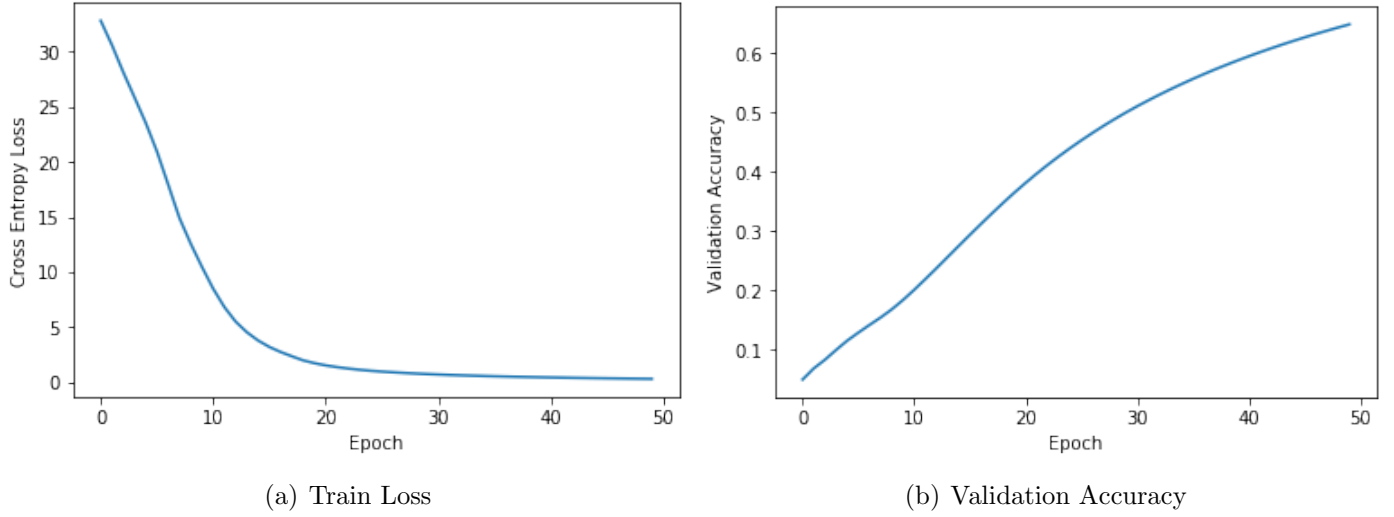


Figure 2.5: Training

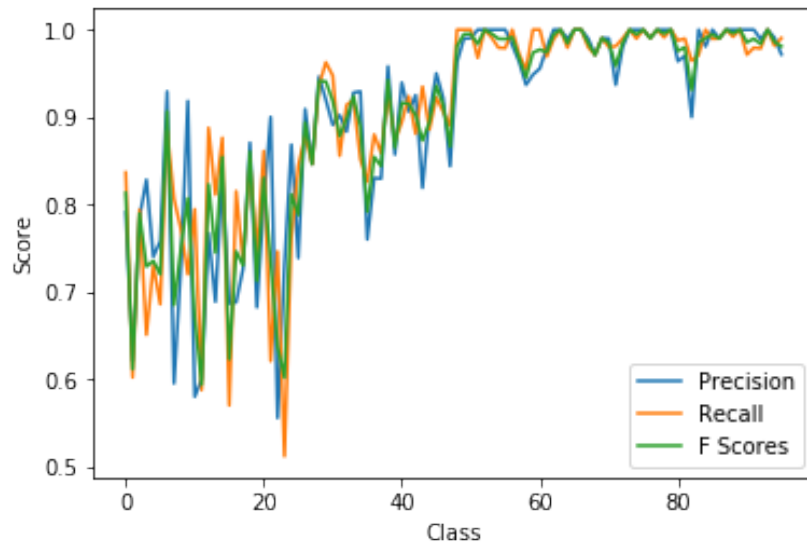


Figure 2.6: Scores

Chapter 3

Inference

- The default parameters of the random normal initializer are 0.0 for mean and 1.0 for the standard deviation. Decreasing the standard deviation to 0.1-0.01 range increased the performance.
- If the model is taking too long to train, then we can try and decrease the batch size to make the updates to weight and bias more frequent. Additionally we can reduce the learning rate as well.
- In the line dataset models, the models are able to identify long lines better than short ones.