

# REPORT

1805061 – Noshin Nawal

## Hyperparameters

1. Batch size = 1024
2. Epoch = 15
3. Default ADAM parameters ( $\beta_1 = 0.9$ ,  $\beta_2 = 0.999$ ,  $\epsilon = 1e-8$ )
4. Xavier initialization
5. Linear learning rate scheduling (initial LR varying from  $5e-3$  to  $5e-5$ , final LR  $5e-6$ )

## Model 1 (Best Model)

```
network = [  
    Dense(28*28, 1024),  
    ReLU(),  
    Dropout(0.3),  
    Dense(1024, 26),  
    Softmax()  
]
```

Initial LR = 5e-3

Test Accuracy = 0.9203

Test F1 = 0.9204

Train Loss = 0.0903

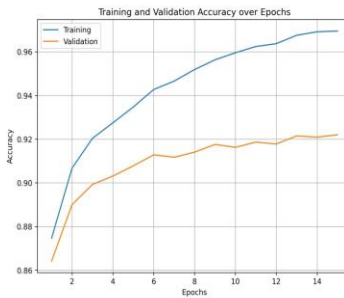
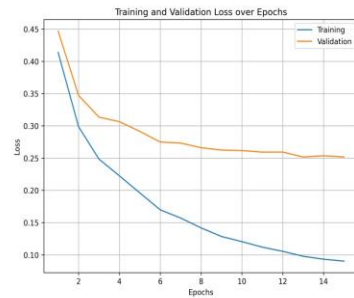
Train Accuracy = 0.9695

Train F1 = 0.9695

Val Loss = 0.2517

Val Accuracy = 0.9220

Val F1 = 0.9217

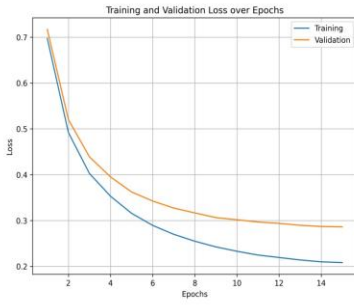


	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	
0	703	1	1	4	1	0	5	13	0	0	2	0	0	7	1	1	23	2	1	2	0	1	0	0	0	4	
1	659	2	1	4	1	6	4	0	0	0	1	0	0	3	1	2	1	0	0	0	0	0	0	0	0	2	
2	1	0	709	0	10	0	1	0	2	0	0	5	0	0	6	0	1	2	1	0	2	0	0	0	0	0	
3	2	7	0	649	0	0	0	2	0	7	0	0	0	1	16	4	2	0	0	0	2	1	0	0	0	1	
4	2	1	15	0	662	1	0	0	0	0	2	0	0	2	1	2	1	0	0	0	0	0	0	1	0	0	
5	1	1	0	0	2	690	0	1	2	1	1	1	0	1	0	9	1	6	3	16	0	0	0	1	0	0	
6	9	6	2	4	0	2	553	0	0	3	1	0	0	1	0	0	78	2	5	1	1	0	1	0	5	3	
7	2	1	0	2	0	2	0	656	3	0	3	5	1	18	0	1	0	1	1	0	1	0	0	0	1	0	
8	0	1	0	0	1	1	0	0	492	18	0	182	0	0	0	0	0	0	2	4	0	0	0	0	0	3	
9	1	0	0	6	0	2	5	0	21	671	0	1	0	0	1	0	0	0	4	6	0	5	1	0	3	0	
10	0	2	4	2	0	0	1	17	0	0	682	3	1	1	0	1	1	5	0	0	0	3	1	11	0	0	
11	0	3	6	0	0	0	0	4	136	2	0	560	0	0	0	0	1	1	0	0	0	0	0	1	0	0	
12	0	0	0	0	0	0	3	0	0	0	0	692	11	0	0	0	0	0	0	1	1	0	2	0	1	0	
13	6	0	0	4	1	0	0	5	0	0	0	9	675	0	0	0	2	0	0	0	0	2	4	1	0	1	
14	3	0	0	11	0	0	0	0	1	0	0	0	1	687	1	1	0	0	0	2	0	0	0	0	0	0	
15	0	0	0	5	0	6	1	0	0	0	0	0	0	1	2	722	4	2	0	0	0	0	0	0	1	0	
16	14	0	1	3	3	2	59	0	0	0	0	0	0	6	6	623	3	0	1	0	0	3	0	2	1	0	
17	4	1	3	1	6	3	0	4	1	0	9	0	3	1	0	4	4	693	0	4	0	6	0	3	2	1	
18	2	4	0	0	0	0	7	0	1	6	0	0	0	0	0	0	2	0	686	0	0	0	0	1	0	0	
19	0	1	1	1	2	10	0	2	6	2	1	5	0	0	0	0	1	8	0	695	0	0	0	0	3	3	
20	2	0	0	1	2	0	1	1	0	2	1	1	1	2	4	0	0	0	0	0	0	693	12	3	0	2	0
21	0	0	0	0	0	0	0	0	1	2	0	0	0	0	0	0	0	5	0	0	0	41	666	1	1	9	0
22	1	1	0	0	0	0	2	1	0	0	1	0	3	8	0	0	0	0	1	0	2	5	705	0	1	0	
23	2	0	0	0	0	0	0	1	2	2	9	0	1	1	0	1	4	1	0	2	0	3	0	654	6	6	
24	0	1	0	0	0	1	3	1	4	2	0	0	1	0	0	2	1	5	0	4	8	14	0	3	675	0	
25	1	4	0	0	3	2	5	0	2	0	0	0	0	0	0	1	2	0	2	3	0	0	0	1	0	708	
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25		

Initial LR = 1e-3

Test Accuracy = 0.9084

Test F1 = 0.9085



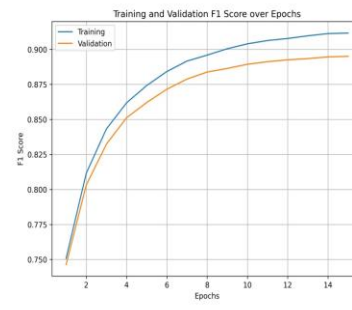
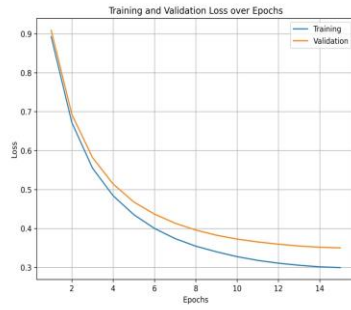
Confusion Matrix

0	683	2	0	6	4	1	4	11	0	0	3	0	2	12	1	0	28	3	1	2	2	0	1	0	1	5
1	3	650	1	4	2	0	7	4	0	0	0	1	0	2	5	3	1	0	0	0	0	0	0	1	0	4
2	1	1	701	1	13	0	2	0	0	0	0	5	0	0	7	0	1	4	2	0	1	0	0	1	0	0
3	2	6	0	644	0	0	1	2	0	10	1	2	0	2	15	5	2	0	0	0	1	1	0	0	0	0
4	3	1	18	0	655	0	1	0	2	0	0	0	0	0	0	2	2	3	1	1	0	0	0	1	0	3
5	0	1	0	0	0	677	1	1	1	1	1	4	0	2	0	17	2	9	3	14	0	0	0	1	1	1
6	14	8	1	4	1	3	532	0	0	0	1	2	0	1	0	0	88	2	11	0	2	0	1	0	4	2
7	5	0	0	2	0	1	0	654	1	0	4	7	1	20	0	0	0	0	0	0	1	0	1	1	0	0
8	0	1	0	1	0	2	0	0	501	16	0	174	0	0	0	0	0	2	3	0	1	0	0	0	0	3
9	0	1	0	9	0	0	5	0	23	663	0	3	0	0	1	0	0	0	8	6	2	4	0	0	2	0
10	1	1	5	3	0	3	1	21	1	0	674	2	2	0	0	1	1	4	0	0	3	3	0	9	0	0
11	0	2	6	1	0	0	0	3	142	2	0	551	0	0	0	0	1	2	1	0	0	0	0	2	1	0
12	1	0	0	0	0	0	1	5	0	0	1	0	687	11	0	0	0	0	0	0	2	1	0	2	0	0
13	5	0	0	5	0	0	0	9	0	0	0	0	5	670	0	0	0	4	0	0	0	1	6	3	1	1
14	3	1	0	11	2	0	2	0	0	1	0	0	0	1	680	1	2	0	0	0	3	0	0	0	0	0
15	3	0	0	4	0	9	1	0	0	0	0	0	1	2	1	712	9	1	0	0	0	0	0	0	1	0
16	13	0	1	4	8	5	55	0	1	1	0	0	0	0	5	6	614	2	1	0	1	0	1	1	4	4
17	9	0	3	1	4	5	0	3	1	0	12	0	4	5	0	3	5	681	0	2	0	6	0	3	3	3
18	2	3	1	1	0	1	9	0	2	6	0	0	0	0	0	0	2	0	681	0	0	0	0	1	0	0
19	0	1	1	1	2	11	1	2	10	2	1	5	0	0	0	0	0	7	0	684	0	0	0	0	10	3
20	3	0	0	2	2	0	0	2	0	2	1	0	1	2	6	0	1	0	0	0	680	18	4	0	3	1
21	1	0	0	0	0	0	0	1	2	1	0	0	2	1	1	0	0	6	0	1	25	672	2	2	9	0
22	1	0	0	1	1	0	0	0	0	0	1	0	2	9	0	0	0	0	1	0	4	5	703	0	3	0
23	3	0	1	0	0	0	0	1	1	2	11	1	1	0	0	0	4	1	0	2	0	3	0	645	11	8
24	0	0	0	0	0	0	5	0	2	2	3	0	1	1	0	2	3	6	0	8	9	22	0	5	655	1
25	1	2	0	2	3	0	8	0	6	0	0	1	1	0	0	0	3	0	2	4	0	0	0	0	0	701
	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25

Initial LR = 5e-4

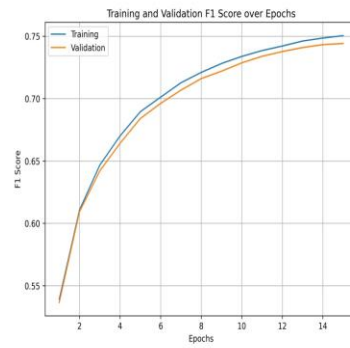
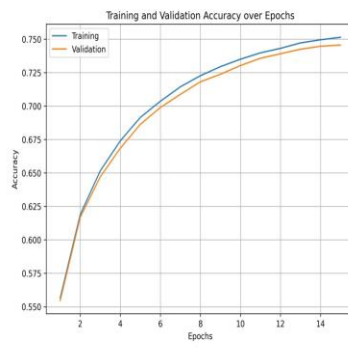
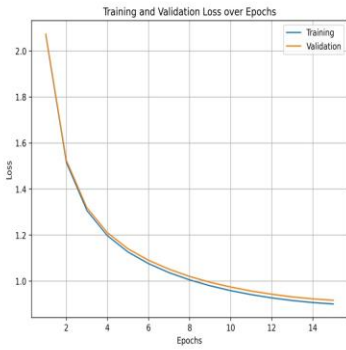
Test Accuracy = 0.8922

Test F1 = 0.8923



		Confusion Matrix																											
True labels	0	664	1	0	9	5	2	6	13	0	0	4	0	4	11	1	1	28	4	1	2	4	0	3	0	2	7		
	1	3	638	0	5	3	0	10	6	0	1	0	2	0	2	7	3	1	0	1	1	0	0	0	0	0	5		
	2	2	1	695	0	15	0	1	0	1	0	1	4	0	0	9	0	2	4	2	0	2	0	0	1	0	0		
	3	5	8	0	632	0	1	1	1	0	11	1	3	0	1	19	4	2	0	2	0	1	1	1	0	0	0		
	4	4	4	4	25	0	630	1	0	0	2	0	0	0	1	0	1	2	4	10	2	1	0	0	0	1	0	5	
	5	0	1	0	1	1	669	1	1	3	2	1	1	0	1	0	22	2	13	2	12	0	0	0	1	2	1		
	6	13	9	4	6	2	3	516	0	0	0	0	2	2	1	2	2	93	2	11	1	2	0	1	0	3	2		
	7	8	0	0	1	1	2	0	641	1	0	4	10	1	23	0	0	0	0	0	0	1	0	2	1	1	1		
	8	0	1	0	1	0	1	0	1	505	12	0	168	0	0	0	0	0	0	2	2	0	0	0	2	4	5		
	9	0	1	1	12	0	0	6	0	26	653	0	3	0	0	1	0	0	0	8	10	1	3	0	0	2	0		
	10	1	4	5	4	0	2	1	26	0	1	657	3	2	0	0	1	0	6	0	0	3	5	1	11	1	1		
	11	0	3	6	2	0	0	0	3	143	3	0	546	0	0	0	0	1	1	0	0	0	0	0	4	2	0		
	12	2	0	0	0	0	0	2	7	0	0	1	0	682	13	0	0	0	0	0	2	1	0	1	0	0	0		
	13	7	1	0	6	0	0	0	9	0	0	1	0	7	654	1	0	0	6	0	0	4	3	9	1	0	1		
	14	4	2	1	8	1	0	2	0	0	1	0	0	0	4	678	1	2	0	0	0	3	0	0	0	0	0		
	15	3	0	0	6	0	10	0	0	0	1	0	0	1	4	1	700	12	2	0	2	0	0	0	0	2	0		
	16	14	1	3	3	11	6	56	0	1	1	1	0	0	0	8	8	601	3	1	2	0	0	0	1	5	1		
	17	11	0	3	0	4	6	0	4	1	0	13	0	4	6	0	7	8	660	0	6	0	6	1	4	5	4		
	18	1	4	1	1	0	1	14	0	3	10	0	0	0	0	0	0	3	0	670	0	0	0	0	1	0	0		
	19	0	0	1	2	4	12	1	6	12	1	2	6	0	0	0	0	1	7	0	671	0	1	0	0	11	3		
	20	7	0	1	3	1	0	0	3	0	2	2	1	1	2	7	0	1	0	0	0	673	20	2	0	1	1		
	21	1	0	0	0	0	0	0	0	1	1	0	1	2	4	1	1	0	7	0	0	26	666	4	3	8	0		
	22	3	0	0	2	1	0	0	1	0	0	2	0	3	11	0	0	0	0	1	1	6	6	692	0	2	0		
	23	3	1	1	0	0	2	1	2	2	2	12	1	2	0	0	0	3	1	0	2	0	5	0	635	12	8		
	24	1	1	0	0	0	1	8	0	4	3	3	0	3	1	0	2	2	6	0	8	8	28	0	6	639	1		
	25	4	2	0	2	4	0	7	0	6	3	0	1	1	0	0	0	4	1	2	5	0	0	1	1	0	690		
	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25			

Test F1 = 0.7462



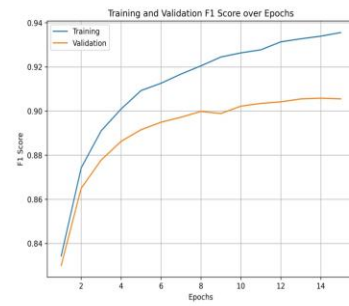
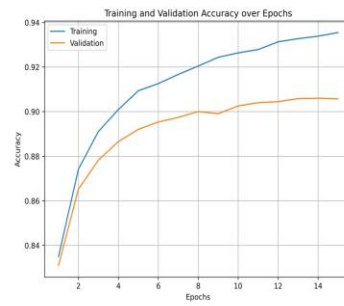
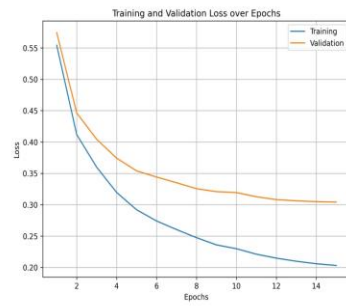
### Confusion Matrix

0	-	491	2	4	30	15	10	16	35	0	3	8	5	18	32	20	4	32	6	0	4	14	0	8	3	2	10	
1	-	6	534	1	18	10	0	26	22	4	6	6	8	4	8	9	2	3	0	2	1	3	0	1	0	4	10	
2	-	9	2	634	0	32	1	7	0	1	0	5	4	0	3	18	2	5	6	1	2	7	0	1	0	0	0	
3	-	5	40	2	492	1	1	1	5	4	16	4	16	0	1	51	8	6	0	5	0	15	7	8	1	0	5	
4	-	23	16	56	2	524	6	6	3	3	0	10	0	1	1	5	2	4	14	3	4	1	1	1	1	0	6	
5	-	3	3	0	1	0	570	3	3	6	3	4	3	2	1	0	47	16	22	3	26	0	2	0	5	11	3	
6	-	23	22	11	6	6	8	363	5	3	16	2	1	7	3	4	15	118	1	29	8	6	2	3	1	10	4	
7	-	12	6	0	13	1	0	0	509	14	0	29	27	3	60	0	3	0	1	0	1	9	0	2	6	1	1	
8	-	0	5	0	2	0	5	0	0	433	28	3	155	1	1	0	1	1	1	3	11	3	3	0	18	8	22	
9	-	1	6	4	34	1	4	6	1	19	570	0	15	1	2	0	1	2	5	14	17	4	6	1	2	4	7	
10	-	10	6	13	2	3	10	0	44	3	1	530	18	5	8	0	1	2	22	0	14	5	14	2	18	4	0	
11	-	1	3	15	4	0	1	0	2	183	4	10	416	0	0	0	2	0	0	0	3	9	12	0	20	23	6	
12	-	8	0	0	0	1	0	1	5	0	0	0	0	0	631	50	1	1	3	0	0	3	3	1	2	0	1	0
13	-	31	2	0	6	0	0	2	20	1	0	9	0	52	493	4	1	6	7	0	4	11	13	34	6	2	6	
14	-	4	2	7	14	1	0	15	1	0	1	0	0	0	15	627	3	8	2	1	1	3	1	1	0	0	0	
15	-	1	0	0	0	0	21	1	0	1	1	1	1	0	10	8	655	12	10	0	12	0	0	4	0	5	1	
16	-	25	2	2	12	11	23	68	1	0	2	2	2	0	1	13	9	486	3	7	23	4	1	5	3	18	4	
17	-	40	1	5	0	12	28	1	15	1	0	42	0	8	15	0	23	23	473	0	22	0	8	6	16	11	3	
18	-	5	7	6	2	2	5	15	0	4	37	3	1	0	0	2	0	12	1	593	1	2	0	0	6	2	3	
19	-	5	7	7	5	16	51	3	20	21	6	10	16	4	1	0	9	4	31	3	450	0	8	0	5	54	5	
20	-	23	1	1	19	2	1	2	8	0	0	9	6	3	5	11	0	0	3	0	0	578	29	23	0	2	2	
21	-	0	1	0	2	1	1	0	2	2	1	6	4	2	17	2	4	1	18	0	3	44	581	7	6	21	0	
22	-	5	6	0	4	1	1	0	12	0	0	3	0	9	10	0	0	0	1	1	1	18	11	646	1	1	0	
23	-	14	1	0	1	0	3	3	5	9	1	33	2	3	0	0	1	4	1	1	4	0	15	2	547	39	6	
24	-	1	3	0	1	0	4	12	6	13	17	3	14	7	2	0	13	2	14	5	38	6	36	3	11	514	0	
25	-	6	9	6	12	5	1	8	1	1	14	5	2	3	9	0	2	8	1	1	8	3	0	2	11	0	616	
		0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	

## Model 2

```
network = [
    Dense(28*28, 256),
    ReLU(),
    Dropout(0.3),
    Dense(256, 26),
    Softmax()
```

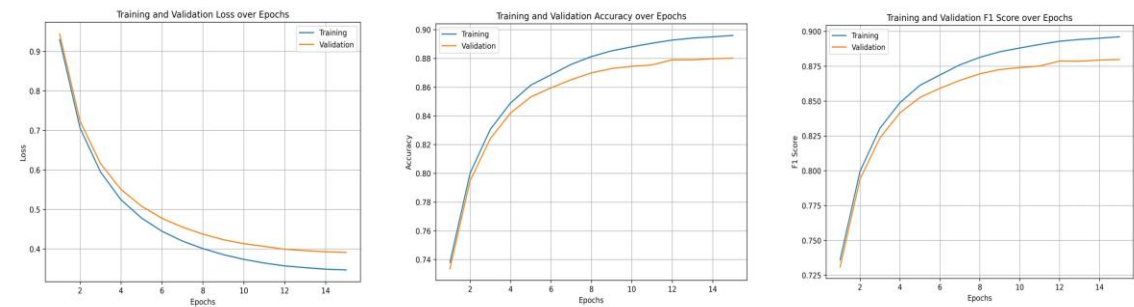
Initial LR = 5e-3



	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
0	673	1	0	9	2	1	5	10	0	0	3	0	3	10	1	1	32	6	0	3	2	1	1	0	0	8
1	1	648	1	4	4	0	4	5	0	1	0	2	1	1	6	3	2	0	0	1	0	0	0	1	0	3
2	0	1	701	0	15	0	1	0	1	0	0	7	0	0	4	0	1	4	2	0	1	0	1	1	0	0
3	3	5	0	647	0	0	0	1	0	9	1	2	0	1	18	5	0	1	0	0	0	1	0	0	0	0
4	3	3	16	0	647	1	1	0	0	0	0	0	1	0	1	2	4	7	1	1	0	0	0	1	0	4
5	3	0	0	1	1	670	2	1	3	1	1	2	0	1	0	20	0	10	2	15	0	0	0	0	3	1
6	14	10	3	3	2	2	517	0	1	1	1	1	0	0	1	2	96	3	10	2	1	0	1	0	4	2
7	4	1	0	3	0	1	0	647	4	0	5	6	3	21	0	0	1	0	0	0	1	0	1	0	0	0
8	0	1	0	1	1	0	0	0	512	14	1	164	0	0	0	0	0	2	3	0	0	0	0	0	2	3
9	1	0	1	11	0	0	4	0	27	657	0	1	0	0	1	0	1	0	6	13	1	3	0	0	0	0
10	1	5	5	2	0	1	0	28	0	0	662	3	2	2	0	0	1	7	0	0	1	3	0	11	1	0
11	0	4	5	2	0	1	0	3	161	2	1	530	0	0	0	0	0	1	1	0	1	1	0	1	0	0
12	2	0	0	0	0	0	0	5	0	0	2	0	681	17	0	0	0	0	0	0	1	1	0	2	0	0
13	7	0	0	4	0	0	0	8	0	0	0	0	11	660	1	0	1	3	0	0	1	2	8	2	1	1
14	2	0	0	11	1	0	1	0	0	1	0	0	0	1	684	1	3	0	0	0	2	0	0	0	0	0
15	2	0	0	4	0	8	1	0	0	0	0	0	1	5	3	710	6	2	0	1	0	0	0	0	1	0
16	15	3	1	3	8	4	47	0	0	0	1	0	0	0	7	7	621	2	0	2	0	0	0	2	3	1
17	7	0	3	0	3	5	1	3	0	1	6	0	2	5	0	6	7	681	0	5	0	7	2	3	3	3
18	1	4	0	1	0	0	11	0	2	4	0	0	0	0	0	0	2	0	684	0	0	0	0	0	0	0
19	0	1	2	0	1	12	0	2	12	4	1	1	0	0	0	1	3	7	0	686	0	0	0	0	4	4
20	4	0	0	1	2	0	0	1	0	3	2	1	2	2	3	0	2	0	0	0	680	19	3	0	2	1
21	0	0	0	0	0	0	0	0	1	1	0	0	1	1	0	0	5	0	0	31	672	4	0	10	0	
22	2	4	0	1	1	0	0	1	0	0	1	0	1	10	0	0	0	0	1	0	8	7	691	1	1	1
23	2	0	1	1	0	0	1	0	1	2	14	0	2	2	0	1	5	2	0	2	0	3	0	639	9	8
24	2	0	0	0	0	0	3	1	2	3	1	2	1	0	0	2	4	6	0	4	9	21	0	7	656	1
25	2	5	0	0	4	0	4	0	4	3	0	1	0	0	0	1	3	0	1	6	0	0	0	1	1	698
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	

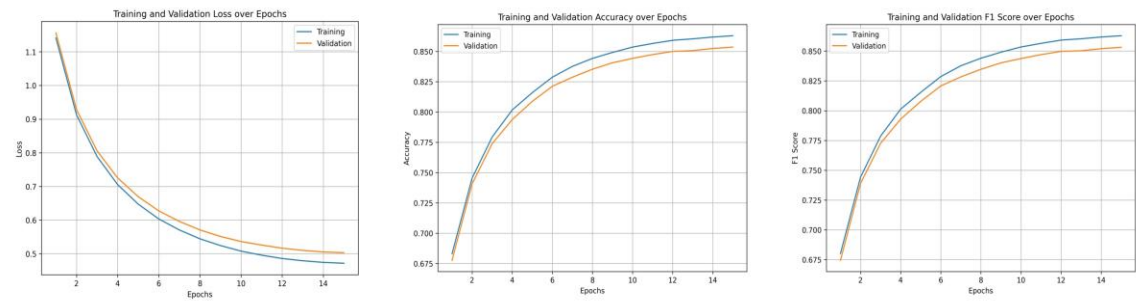


Initial LR = 1e-3



Confusion Matrix	
True labels	Predicted labels
0	650 3 2 12 6 4 9 13 0 0 4 0 6 12 3 0 27 2 0 2 6 1 3 0 0 7
1	4 623 1 9 6 1 11 7 0 1 1 3 1 4 7 1 1 0 0 0 0 0 1 0 1 5
2	2 1 686 0 22 0 3 1 1 0 1 2 0 1 9 0 2 5 2 0 1 0 0 1 0 0
3	7 20 0 609 0 2 0 4 0 13 3 3 0 1 20 5 1 0 1 0 2 2 0 0 0 1
4	5 2 32 0 616 1 1 0 2 0 1 0 0 0 5 3 5 11 2 2 0 0 0 1 0 4
5	0 1 0 1 2 655 1 1 4 3 0 2 0 2 0 26 2 14 2 15 0 0 0 1 4 1
6	12 7 5 6 3 1 495 0 0 2 1 2 5 2 3 1 106 1 9 1 2 0 2 1 6 4
7	10 2 0 4 1 2 0 621 3 0 6 9 2 29 0 0 0 0 0 0 1 0 3 4 0 1
8	0 1 0 3 0 1 0 2 489 14 0 172 0 0 0 0 0 0 0 4 0 1 0 3 4 10
9	1 1 2 19 0 0 5 0 26 644 0 4 0 1 0 0 0 0 6 13 1 2 0 0 2 0
10	2 4 4 4 1 1 0 23 0 0 656 1 2 1 0 1 1 11 0 1 4 4 1 12 1 0
11	0 3 8 3 0 0 0 0 156 3 0 532 0 0 0 0 1 1 0 0 1 0 0 3 2 1
12	5 0 0 0 0 0 2 9 0 0 2 0 671 16 0 0 0 0 0 2 1 0 2 1 0 0
13	12 0 0 5 0 0 0 13 0 0 1 0 10 640 1 2 0 5 0 1 6 2 8 1 1 2
14	2 2 1 10 1 0 2 0 0 0 0 0 0 3 679 1 3 0 0 0 3 0 0 0 0 0
15	3 0 0 6 0 8 4 0 0 1 0 0 2 6 2 695 7 7 0 2 0 0 0 0 1 0
16	20 0 1 4 11 6 56 0 1 0 1 0 1 1 10 8 589 4 0 4 1 0 1 1 6 1
17	16 0 2 0 5 4 1 6 1 0 13 0 3 11 0 8 5 650 1 7 0 7 1 5 3 4
18	2 4 1 3 0 2 12 0 2 12 0 0 0 0 0 0 4 0 665 0 1 0 0 1 0 0
19	1 1 2 4 5 15 1 4 11 1 4 5 0 0 0 0 1 8 1 662 0 0 0 1 11 3
20	8 0 1 7 1 0 0 2 0 2 3 1 1 3 7 0 1 0 0 0 665 21 2 0 2 1
21	1 0 0 1 0 0 0 0 1 1 0 1 2 3 1 0 1 6 0 0 37 653 3 3 12 0
22	3 0 0 3 1 0 0 1 0 0 3 0 3 9 0 0 0 0 1 2 7 7 691 0 0 0
23	4 2 1 1 0 2 1 1 4 1 16 1 2 1 0 1 3 1 0 2 0 4 2 627 12 6
24	1 2 0 0 0 0 9 0 5 6 3 0 4 0 0 2 3 9 1 7 8 31 0 6 627 1
25	2 4 0 3 6 1 7 0 4 3 1 1 2 0 1 1 2 1 2 5 0 0 1 0 0 687

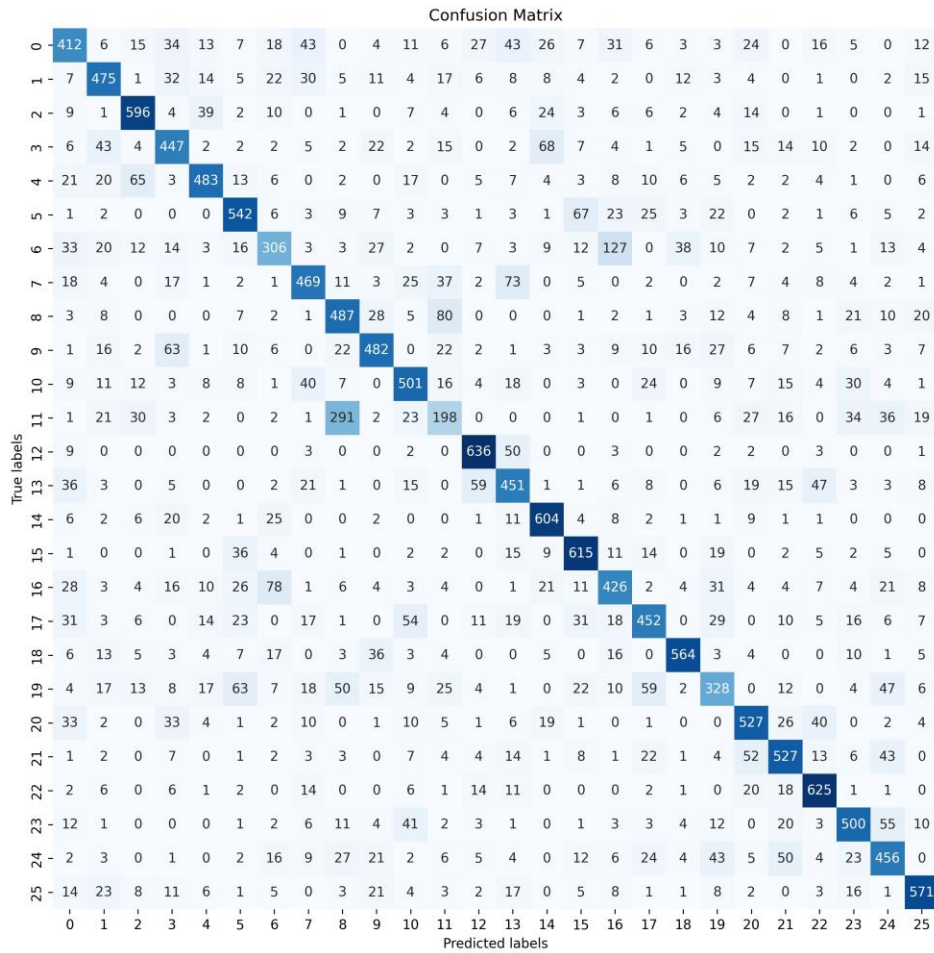
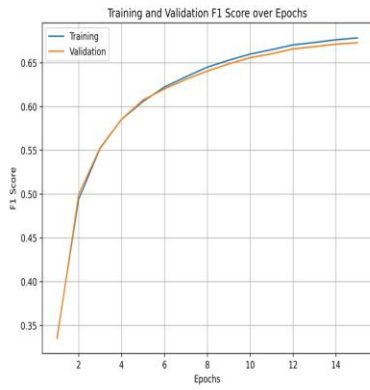
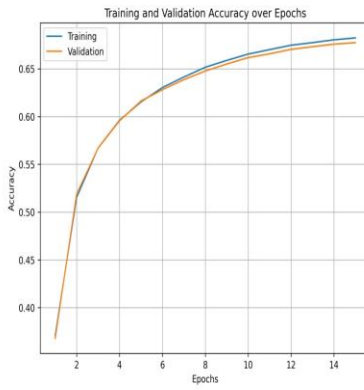
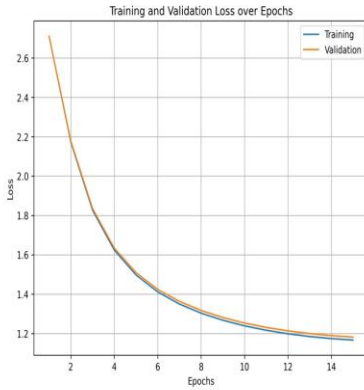
Initial LR = 5e-4



Confusion Matrix	
True labels	0
	1
	2
	3
	4
	5
	6
	7
	8
	9
	10
	11
	12
	13
	14
	15
	16
	17
	18
	19
	20
	21
	22
	23
	24
	25
Predicted labels	



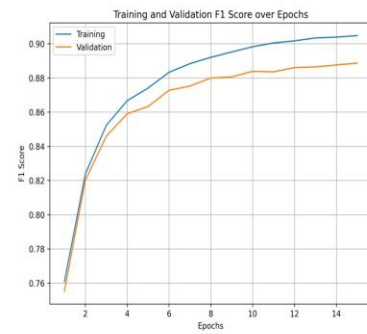
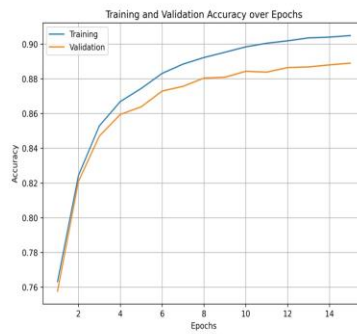
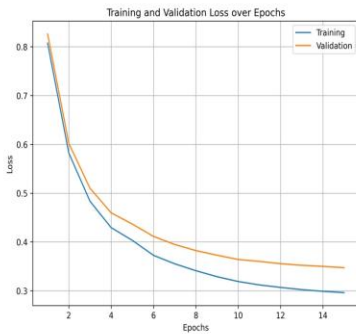
**Initial LR = 5e-5**



## Model 3

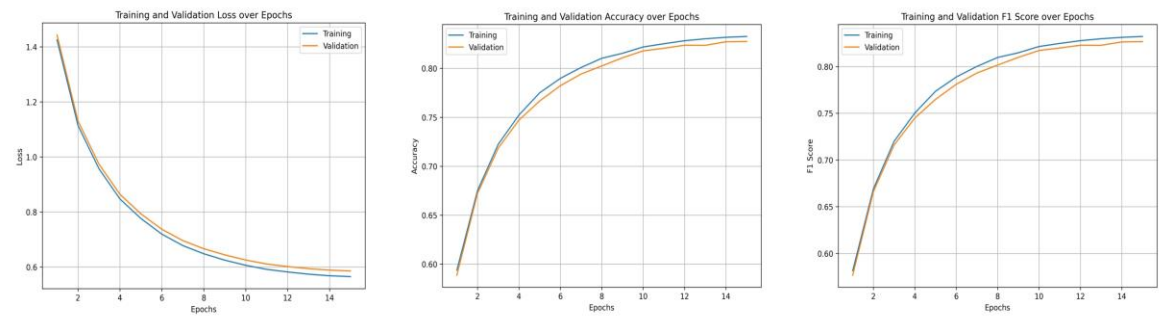
```
network = [
    Dense(28*28, 128),
    ReLU(),
    Dropout(0.2),
    Dense(128, 64),
    ReLU(),
    Dropout(0.1),
    Dense(64, 26),
    Softmax()]
```

Initial LR = 5e-3



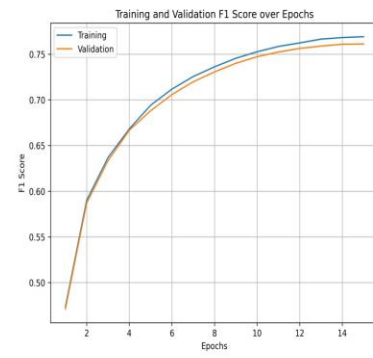
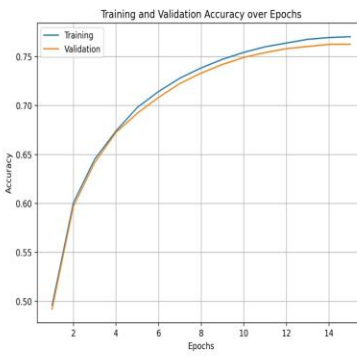
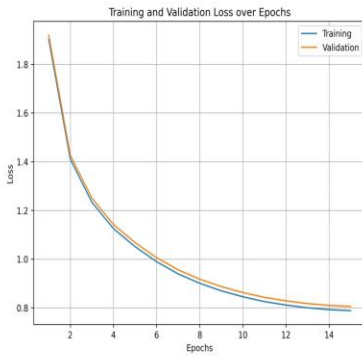
		Confusion Matrix																									
True labels	0	678	2	1	6	9	1	4	10	0	0	2	1	4	6	4	0	24	9	0	2	1	0	1	0	0	7
	1	5	632	0	8	4	0	12	7	0	2	0	1	0	4	8	2	1	0	0	0	0	0	0	0	0	2
	2	7	0	684	0	22	0	0	0	0	0	0	7	0	0	12	0	0	3	3	0	2	0	0	0	0	0
	3	3	5	0	627	0	0	0	1	1	11	3	0	0	0	28	6	3	0	1	0	1	1	1	1	0	2
	4	3	2	24	0	631	1	0	0	0	0	0	0	0	0	2	1	7	9	4	1	0	0	0	1	0	7
	5	0	1	0	0	0	669	0	1	4	1	0	1	1	1	0	25	1	10	2	18	0	0	0	0	1	1
	6	16	5	1	4	1	2	493	2	0	0	1	2	1	0	2	1	117	1	13	2	2	0	2	2	4	3
	7	9	5	0	4	0	1	0	634	1	1	5	9	3	22	0	0	0	2	0	0	1	0	1	0	0	0
	8	1	1	0	2	0	0	0	0	491	13	1	182	0	0	0	0	0	1	2	2	0	1	0	0	0	7
	9	0	0	0	1	10	0	0	5	0	27	647	0	1	0	0	1	0	0	10	12	2	8	0	0	3	0
	10	2	2	2	2	1	0	1	33	0	1	654	2	2	3	0	0	0	8	0	3	2	3	1	13	0	0
	11	0	6	6	2	0	0	0	3	141	2	0	543	0	0	0	0	1	2	0	1	2	0	0	3	2	0
	12	4	0	0	0	0	0	0	10	0	0	1	0	669	17	1	0	0	0	0	1	3	0	5	0	0	0
	13	10	0	0	4	0	0	0	8	0	0	3	0	18	641	1	1	2	1	0	0	3	3	13	1	0	1
	14	2	1	3	10	1	0	2	0	0	1	0	0	0	2	679	1	2	0	0	0	3	0	0	0	0	0
	15	2	0	0	0	5	0	7	1	0	0	0	0	1	1	2	1	706	12	3	0	0	0	0	2	0	1
	16	26	3	1	1	7	10	59	1	0	2	0	0	1	0	12	6	583	3	1	3	0	0	1	0	6	1
	17	17	1	1	0	9	7	1	7	1	0	13	0	4	8	1	6	4	643	0	10	0	9	2	3	3	3
	18	2	3	0	1	0	1	10	0	2	9	0	0	0	0	1	0	2	0	677	0	0	0	0	1	0	0
	19	0	1	2	3	4	11	1	2	9	1	1	6	0	0	0	0	3	10	0	676	0	0	1	0	7	3
	20	8	0	1	3	1	0	0	3	0	2	0	0	1	5	5	0	1	0	0	0	666	22	6	0	3	1
	21	0	0	0	0	0	1	0	0	1	1	0	1	2	1	0	0	1	4	0	1	28	669	3	2	11	0
	22	2	1	0	1	2	0	0	0	0	0	4	0	5	7	0	0	0	0	1	0	8	9	689	0	2	0
	23	6	1	0	0	0	2	3	2	2	2	16	1	2	2	0	0	2	0	0	1	1	4	0	629	12	7
	24	1	0	0	1	0	0	6	1	2	5	0	0	2	0	0	3	2	7	1	10	8	34	0	3	639	0
	25	2	3	1	1	6	0	6	0	5	2	2	0	1	0	0	0	4	0	1	4	0	0	1	1	1	693
	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	

Initial LR = 1e-3



		Confusion Matrix																									
True labels	0	587	6	3	16	11	4	8	24	0	0	4	2	10	15	19	2	23	8	1	2	11	0	2	3	0	11
	1	5	595	0	8	6	2	19	16	0	2	1	3	0	7	8	1	1	0	0	0	1	0	0	0	1	12
	2	2	2	664	0	32	1	4	0	1	0	1	4	0	2	16	0	3	3	2	0	2	0	0	1	0	0
	3	6	14	4	568	2	1	1	3	1	16	4	7	0	3	34	7	3	0	1	0	11	2	4	0	1	1
	4	8	7	44	0	579	5	2	0	2	0	2	0	1	2	8	2	5	15	3	1	0	0	0	1	0	6
	5	4	1	0	1	0	615	2	1	5	4	1	3	0	1	0	47	9	13	3	23	0	0	0	0	3	1
	6	18	13	6	6	4	1	432	1	1	8	2	1	4	1	5	12	119	0	23	4	3	0	3	0	6	4
	7	9	7	0	4	1	0	0	596	3	1	12	15	3	37	0	1	0	2	0	0	3	1	1	0	1	1
	8	0	3	0	2	0	2	0	1	465	22	1	176	0	0	0	0	0	1	4	4	0	0	0	4	7	12
	9	1	1	2	31	1	2	5	1	24	613	0	6	0	0	1	0	0	2	12	10	3	4	0	0	4	4
	10	5	4	7	2	1	4	0	40	0	0	609	5	2	8	0	0	1	11	0	10	2	6	2	13	2	1
	11	0	2	9	2	0	0	0	7	152	2	8	502	0	0	0	0	0	1	0	0	5	0	0	9	7	8
	12	10	0	0	0	0	0	3	8	0	0	1	0	642	36	2	1	1	1	0	1	3	0	2	0	0	0
	13	17	0	0	9	1	0	1	16	1	0	8	0	33	574	4	0	1	5	0	0	11	4	21	1	1	2
	14	5	2	6	10	2	0	3	0	0	1	0	0	0	5	665	1	5	0	0	0	2	0	0	0	0	0
	15	3	0	0	4	1	12	0	0	1	0	0	1	1	8	7	683	11	4	0	4	0	0	1	0	3	0
	16	28	2	5	8	13	13	82	1	1	2	0	1	0	1	16	11	513	1	5	6	0	0	4	1	12	1
	17	27	0	1	0	12	12	2	8	1	0	33	1	6	14	0	14	5	561	0	17	0	7	4	9	14	5
	18	2	5	5	4	0	2	11	0	2	18	0	0	0	0	1	0	9	0	647	1	1	0	0	1	0	0
	19	3	2	6	6	13	21	2	3	13	4	3	11	1	0	0	3	4	18	4	600	0	1	0	0	20	3
	20	11	0	1	11	1	0	0	3	0	0	5	1	3	4	8	0	0	1	0	1	644	20	12	0	2	0
	21	1	0	0	2	1	1	0	3	0	1	1	2	1	5	0	0	1	7	0	1	40	636	3	3	17	0
	22	4	3	0	4	0	1	0	8	0	0	6	0	9	16	0	0	1	2	1	2	12	7	655	0	0	0
	23	6	1	0	0	0	2	0	7	4	2	29	1	2	1	0	1	3	0	0	3	0	10	2	593	24	4
	24	3	2	0	3	0	5	10	0	2	6	2	5	4	1	0	3	6	15	1	16	7	32	1	5	596	0
	25	5	12	4	4	8	2	5	0	3	8	2	3	1	1	0	2	6	1	2	2	0	1	1	5	1	655
		0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25

**Initial LR = 5e-4**



	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
0	509	6	8	30	9	5	14	29	0	1	8	2	18	29	30	2	29	6	0	3	16	0	4	0	1	13
1	7	544	1	16	9	2	22	32	0	4	2	5	2	5	10	1	2	0	5	2	3	0	1	0	1	12
2	6	1	638	2	36	1	5	0	1	1	3	6	0	2	19	1	3	6	2	1	2	0	1	0	0	3
3	7	26	6	505	2	0	3	5	2	18	4	13	0	3	48	7	4	1	3	0	17	5	6	0	1	8
4	14	13	58	1	539	5	3	0	2	1	7	0	3	0	6	6	5	12	4	5	1	0	1	0	0	7
5	5	2	0	1	0	578	2	1	8	7	1	2	0	0	0	58	9	23	2	27	0	2	0	3	4	2
6	27	17	12	5	1	8	369	2	3	20	1	0	5	6	5	14	116	1	35	4	4	2	4	1	12	3
7	12	11	0	6	1	0	1	545	4	0	18	23	5	47	0	2	0	1	0	1	9	1	7	3	0	1
8	0	4	0	1	0	3	0	2	467	25	3	140	1	0	0	0	0	1	3	8	3	0	0	12	5	26
9	1	2	3	48	1	3	5	1	20	574	0	12	0	1	0	0	1	6	13	14	5	3	0	1	7	6
10	7	5	10	3	4	4	0	43	2	0	555	13	2	13	0	1	1	28	0	12	2	7	3	16	3	1
11	0	0	15	2	0	1	0	2	209	3	11	411	0	0	0	0	0	1	0	2	6	2	0	14	16	19
12	9	0	0	0	0	0	2	7	0	0	0	0	0	629	49	3	1	2	1	0	1	3	1	3	0	0
13	20	0	0	9	1	0	2	27	1	0	10	0	47	507	7	1	2	7	0	2	9	12	39	1	3	3
14	8	1	6	15	1	0	7	0	0	1	0	0	0	7	643	2	7	1	1	1	5	0	0	0	1	0
15	4	0	0	2	1	29	0	0	0	0	1	1	0	11	8	652	8	12	0	9	0	0	2	0	4	0
16	41	3	7	10	13	24	87	1	1	2	0	3	0	2	17	7	457	3	6	16	2	0	4	2	17	2
17	41	0	3	0	14	20	1	11	1	0	50	1	8	16	0	24	10	489	0	22	0	8	5	11	16	2
18	5	9	4	5	2	2	13	0	3	23	0	0	0	0	2	0	11	0	623	1	2	0	0	3	0	1
19	4	4	8	6	17	30	5	16	15	10	7	14	3	0	0	9	2	35	4	499	0	4	2	1	43	3
20	10	0	2	19	1	0	0	4	0	1	8	1	2	8	11	0	0	2	0	0	599	36	23	1	0	0
21	0	0	0	5	0	2	0	2	0	0	4	4	4	12	1	1	2	13	0	1	44	595	7	3	26	0
22	4	5	1	4	1	0	0	13	0	0	5	0	11	18	0	0	0	2	1	2	15	7	641	1	0	0
23	12	0	0	0	0	5	2	6	7	1	39	3	2	0	0	1	2	2	0	3	0	18	1	550	34	7
24	2	5	0	2	0	4	11	5	4	13	2	6	6	4	0	3	5	19	2	46	6	38	2	7	533	0
25	5	17	4	4	5	2	7	0	2	9	3	11	3	6	0	2	5	2	4	3	2	0	1	10	2	625
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	

Initial LR = 5e-5

