CSE 472 : Machine Learning Sessional
Offline 3 Report

Sabah Ahmed
1905118

November 7, 2024

1 Instructions for Running Code

Below the cell named **Hyperparameters, Model and File Paths Initialization**, there is a cell for tuning various things and it looks like this.

```
batch_size = 64
num_epochs = 500

# learning rate
lr = 0.01

# depends on how many models are defined
# in my case there 3 models, so possible values are 0, 1, 2
model_id = 0

model_path = f'model_{model_id}_{lr}.pickle'
metrics_base_path = f'metrics_{model_id}_{lr}'
confusion_matrix_path = f'confusion_matrix_{model_id}_{lr}'
```

Here we can

- Change `batch_size` for training (try to keep it a divisor of training examples count).
- Change the number of **epochs** by changing `num_epochs`.
- Change the **learning rate** by changing `lr`.
- Choose any specific **model** by changing `model_id`.

A new model can be created and added to the model list for comparison.

To start the training loop, save the metrics and plot them, some codes need to be uncommented which are clearly mentioned in the notebook.

2 Models's Performance Metrics and Plots Analysis

2.1 Model 0 (Simple Model without BatchNorm and Dropout)

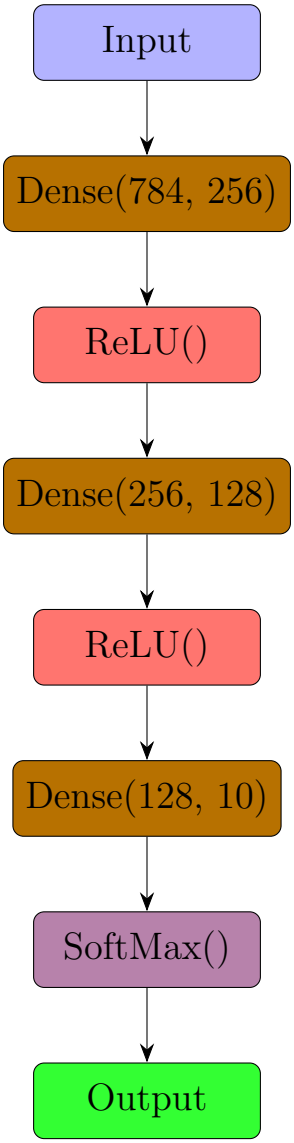

Figure 1: **Model 0 Architecture**

2.1.1 Learning Rate 0.0005

Performance Metrics Table

Epoch	Train Loss	Train Accuracy	Validation Loss	Validation Accuracy	Validation Macro-F1
10	0.231	0.9216	0.3027	0.8932	0.894
20	0.1507	0.9494	0.3157	0.8965	0.8968
30	0.1014	0.9671	0.3558	0.8959	0.8955
40	0.0667	0.9799	0.4087	0.8955	0.8948
50	0.0424	0.985	0.4492	0.8961	0.8958
60	0.0267	0.9935	0.4909	0.8971	0.8969
70	0.017	0.9968	0.536	0.8952	0.895
80	0.0115	0.998	0.5925	0.8973	0.8976
90	0.0067	0.9995	0.6362	0.8952	0.8951
100	0.0045	0.9999	0.6679	0.8972	0.8972
110	0.0025	1.0	0.7139	0.8958	0.8957
120	0.0021	1.0	0.7405	0.8949	0.8951
130	0.0011	1.0	0.7621	0.8952	0.8954
140	0.0007	1.0	0.7917	0.8944	0.8946
150	0.0005	1.0	0.8224	0.8948	0.8948
160	0.0004	1.0	0.8414	0.8943	0.8945
170	0.0003	1.0	0.8578	0.8946	0.8946
180	0.0002	1.0	0.8711	0.8944	0.8946
190	0.0002	1.0	0.8889	0.8944	0.8944
200	0.0002	1.0	0.8963	0.8944	0.8944
210	0.0001	1.0	0.9063	0.8948	0.8948
220	0.0001	1.0	0.9131	0.8942	0.8942
230	0.0001	1.0	0.9194	0.8946	0.8946
240	0.0001	1.0	0.9254	0.8948	0.8948
250	0.0001	1.0	0.93	0.8946	0.8945
260	0.0001	1.0	0.9322	0.8945	0.8945
270	0.0001	1.0	0.934	0.8941	0.8941
280	0.0001	1.0	0.937	0.8944	0.8944
290	0.0001	1.0	0.939	0.8945	0.8945
300	0.0001	1.0	0.9411	0.8942	0.8942
310	0.0001	1.0	0.9422	0.8945	0.8945
320	0.0001	1.0	0.943	0.8944	0.8944
330	0.0001	1.0	0.9444	0.8945	0.8945
340	0.0001	1.0	0.9448	0.8942	0.8943
350	0.0001	1.0	0.9456	0.8943	0.8943
360	0.0001	1.0	0.9458	0.8944	0.8944
370	0.0001	1.0	0.9463	0.8945	0.8945
380	0.0001	1.0	0.9466	0.8943	0.8943
390	0.0001	1.0	0.9469	0.8944	0.8944
400	0.0001	1.0	0.9471	0.8943	0.8943
410	0.0001	1.0	0.9472	0.8945	0.8945
420	0.0001	1.0	0.9474	0.8944	0.8944
430	0.0001	1.0	0.9475	0.8945	0.8945
440	0.0001	1.0	0.9476	0.8945	0.8945
450	0.0001	1.0	0.9477	0.8944	0.8944
460	0.0001	1.0	0.9478	0.8944	0.8944
470	0.0001	1.0	0.9478	0.8944	0.8944
480	0.0001	1.0	0.9479	0.8944	0.8944
490	0.0001	1.0	0.9479	0.8944	0.8944
500	0.0001	1.0	0.9479	0.8944	0.8944

Plots

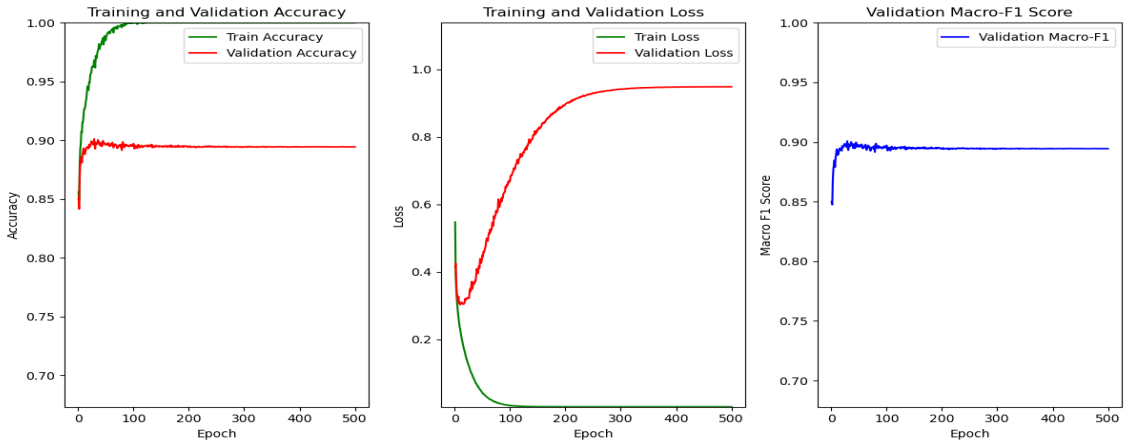

Figure 2: Various Plots

Confusion Matrix on Validation Data

Actual \ Predicted	Class 0	Class 1	Class 2	Class 3	Class 4	Class 5	Class 6	Class 7	Class 8	Class 9
Class 0	1051	0	21	39	9	1	77	0	4	0
Class 1	1	1189	2	18	5	0	2	0	1	1
Class 2	25	0	1016	10	90	0	60	0	4	0
Class 3	27	6	11	1101	20	0	13	0	6	0
Class 4	6	0	101	42	1009	0	38	0	6	0
Class 5	0	0	0	1	0	1176	0	19	1	14
Class 6	147	2	88	39	78	0	856	0	7	1
Class 7	0	0	0	0	0	19	0	1104	4	32
Class 8	10	0	7	4	6	6	6	4	1153	1
Class 9	0	0	0	0	0	15	0	28	1	1159

Table 1: Confusion Matrix

2.1.2 Learning Rate 0.001

Performance Metrics Table

Epoch	Train Loss	Train Accuracy	Validation Loss	Validation Accuracy	Validation Macro-F1
10	0.2218	0.9207	0.3102	0.892	0.8932
20	0.1415	0.9541	0.3416	0.8955	0.8959
30	0.0872	0.9677	0.413	0.8914	0.8915
40	0.0567	0.9857	0.4764	0.8968	0.8969
50	0.032	0.9871	0.5849	0.8925	0.892
60	0.0187	0.9941	0.6479	0.8938	0.8942
70	0.0097	0.9955	0.694	0.8958	0.8958
80	0.0059	0.9988	0.7474	0.8938	0.894
90	0.0019	0.9994	0.7819	0.8949	0.8953
100	0.001	1.0	0.8163	0.8988	0.8989
110	0.0028	1.0	0.839	0.8949	0.895
120	0.0015	1.0	0.8637	0.8962	0.8964
130	0.0003	1.0	0.8982	0.8973	0.8974
140	0.0009	1.0	0.9209	0.8959	0.896
150	0.0001	1.0	0.9434	0.8964	0.8964
160	0.0001	1.0	0.9746	0.8954	0.8956
170	0.0	1.0	0.9926	0.8966	0.8966
180	0.0	1.0	1.0184	0.8957	0.8957
190	0.0	1.0	1.0382	0.8953	0.8955
200	0.0	1.0	1.0556	0.8958	0.8958
210	0.0	1.0	1.0697	0.8961	0.8962
220	0.0	1.0	1.0826	0.8962	0.8963
230	0.0	1.0	1.093	0.8955	0.8956
240	0.0	1.0	1.1014	0.8962	0.8963
250	0.0	1.0	1.1083	0.8958	0.8959
260	0.0	1.0	1.1136	0.8955	0.8956
270	0.0	1.0	1.1188	0.896	0.8961
280	0.0	1.0	1.1223	0.8955	0.8956
290	0.0	1.0	1.1253	0.8958	0.8958
300	0.0	1.0	1.1282	0.8956	0.8957
310	0.0	1.0	1.1306	0.8952	0.8953
320	0.0	1.0	1.1319	0.8955	0.8956
330	0.0	1.0	1.1337	0.8954	0.8955
340	0.0	1.0	1.1346	0.8955	0.8956
350	0.0	1.0	1.1359	0.8955	0.8956
360	0.0	1.0	1.1365	0.8953	0.8954
370	0.0	1.0	1.1371	0.8953	0.8954
380	0.0	1.0	1.1377	0.8955	0.8956
390	0.0	1.0	1.1382	0.8956	0.8957
400	0.0	1.0	1.1386	0.8955	0.8956
410	0.0	1.0	1.1388	0.8955	0.8956
420	0.0	1.0	1.1391	0.8954	0.8955
430	0.0	1.0	1.1392	0.8954	0.8955
440	0.0	1.0	1.1394	0.8954	0.8955
450	0.0	1.0	1.1395	0.8954	0.8955
460	0.0	1.0	1.1396	0.8954	0.8955
470	0.0	1.0	1.1397	0.8954	0.8955
480	0.0	1.0	1.1398	0.8954	0.8955
490	0.0	1.0	1.1399	0.8954	0.8955
500	0.0	1.0	1.1399	0.8954	0.8955

Plots

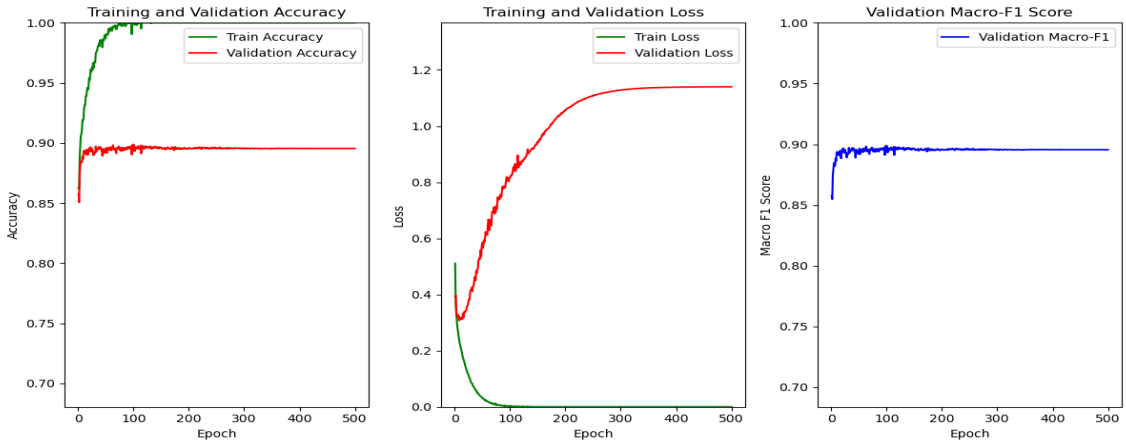

Figure 3: Various Plots

Confusion Matrix on Validation Data

Actual \ Predicted	Class 0	Class 1	Class 2	Class 3	Class 4	Class 5	Class 6	Class 7	Class 8	Class 9
Class 0	1025	2	28	41	5	0	98	0	3	0
Class 1	0	1190	1	19	6	0	2	0	1	0
Class 2	29	0	1030	7	77	1	59	0	2	0
Class 3	31	6	19	1071	31	0	19	0	7	0
Class 4	4	1	84	41	1001	0	62	2	7	0
Class 5	0	1	1	0	0	1172	1	20	2	14
Class 6	119	5	101	30	62	0	891	1	9	0
Class 7	0	0	0	0	0	21	1	1099	4	34
Class 8	3	0	6	3	7	7	14	5	1150	2
Class 9	0	0	0	0	0	11	0	33	2	1157

Table 2: Confusion Matrix

2.1.3 Learning Rate 0.005

Performance Metrics Table

Epoch	Train Loss	Train Accuracy	Validation Loss	Validation Accuracy	Validation Macro-F1
10	0.2695	0.9072	0.3401	0.8831	0.8842
20	0.2058	0.9332	0.3695	0.8879	0.8879
30	0.1655	0.9404	0.4327	0.8852	0.8843
40	0.1289	0.9479	0.5075	0.8839	0.8828
50	0.1084	0.9625	0.59	0.8869	0.8865
60	0.0867	0.9709	0.6028	0.8878	0.8882
70	0.068	0.9746	0.7201	0.8873	0.8876
80	0.0548	0.9809	0.7498	0.8888	0.8892
90	0.0442	0.9818	0.8406	0.8872	0.8873
100	0.0347	0.9893	0.8863	0.8892	0.8896
110	0.0269	0.9892	0.9825	0.8858	0.8857
120	0.0217	0.9927	1.0323	0.8878	0.8883
130	0.0166	0.9951	1.0688	0.8858	0.8862
140	0.0146	0.9959	1.1181	0.8857	0.8859
150	0.0117	0.9972	1.1518	0.887	0.8873
160	0.01	0.9975	1.1942	0.8848	0.8851
170	0.0085	0.9979	1.218	0.8852	0.8855
180	0.0072	0.998	1.2461	0.8843	0.8846
190	0.0065	0.9981	1.2635	0.8845	0.8847
200	0.0058	0.9983	1.2801	0.8842	0.8845
210	0.0054	0.9984	1.2958	0.8845	0.8849
220	0.005	0.9986	1.3126	0.8852	0.8856
230	0.0047	0.9987	1.3197	0.8843	0.8846
240	0.0045	0.9987	1.3274	0.8842	0.8845
250	0.0043	0.9989	1.3381	0.8847	0.885
260	0.0041	0.9989	1.3415	0.8844	0.8847
270	0.004	0.9989	1.3469	0.8839	0.8843
280	0.0039	0.9989	1.3509	0.8838	0.8842
290	0.0038	0.9989	1.3547	0.8843	0.8846
300	0.0037	0.9989	1.3573	0.8842	0.8846
310	0.0037	0.9989	1.3609	0.8842	0.8845
320	0.0036	0.9989	1.3627	0.8845	0.8848
330	0.0036	0.9989	1.3641	0.8845	0.8848
340	0.0035	0.9989	1.365	0.8847	0.885
350	0.0035	0.999	1.3661	0.8844	0.8847
360	0.0035	0.9989	1.3671	0.8844	0.8848
370	0.0035	0.999	1.3678	0.8845	0.8848
380	0.0035	0.999	1.3684	0.8847	0.885
390	0.0035	0.999	1.369	0.8847	0.885
400	0.0035	0.999	1.3693	0.8848	0.8851
410	0.0034	0.999	1.3696	0.8848	0.8851
420	0.0034	0.999	1.3699	0.8846	0.8849
430	0.0034	0.999	1.3701	0.8848	0.8851
440	0.0034	0.999	1.3703	0.8846	0.8849
450	0.0034	0.999	1.3704	0.8848	0.8851
460	0.0034	0.999	1.3705	0.8847	0.885
470	0.0034	0.999	1.3706	0.8848	0.8851
480	0.0034	0.999	1.3707	0.8847	0.885
490	0.0034	0.999	1.3708	0.8847	0.885
500	0.0034	0.999	1.3708	0.8847	0.885

Plots

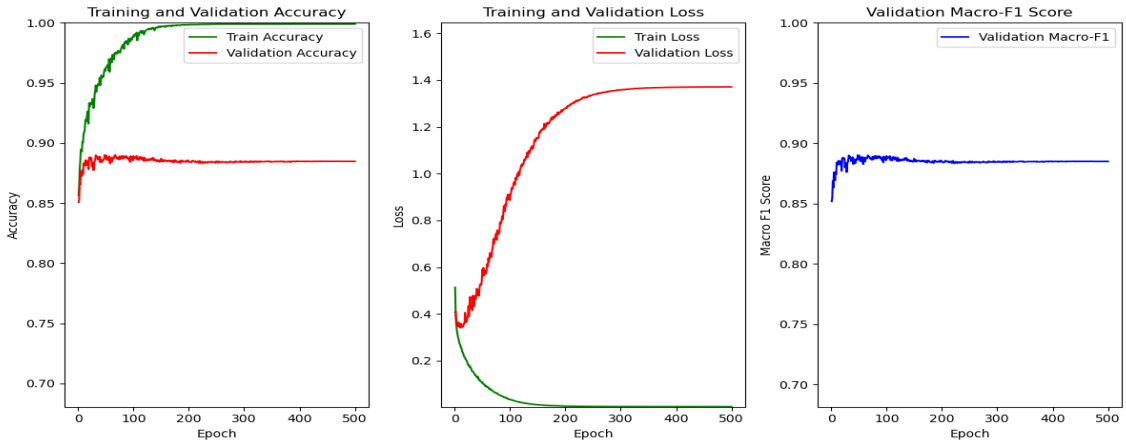

Figure 4: Various Plots

Confusion Matrix on Validation Data

Actual \ Predicted	Class 0	Class 1	Class 2	Class 3	Class 4	Class 5	Class 6	Class 7	Class 8	Class 9
Class 0	999	3	22	46	3	0	127	0	2	0
Class 1	1	1188	1	21	3	0	4	0	1	0
Class 2	19	1	949	11	123	1	99	0	2	0
Class 3	22	6	13	1085	29	1	26	0	2	0
Class 4	1	1	56	49	1018	0	70	0	7	0
Class 5	0	0	0	0	0	1175	1	25	2	8
Class 6	135	3	70	44	78	0	882	0	6	0
Class 7	0	0	0	0	0	17	0	1098	1	43
Class 8	6	0	7	4	8	4	19	7	1141	1
Class 9	0	0	0	0	0	17	0	40	1	1145

Table 3: Confusion Matrix

2.1.4 Learning Rate 0.01

Performance Metrics Table

Epoch	Train Loss	Train Accuracy	Validation Loss	Validation Accuracy	Validation Macro-F1
10	0.3186	0.8918	0.382	0.8727	0.8736
20	0.264	0.9096	0.3895	0.8789	0.8782
30	0.2291	0.9176	0.411	0.8784	0.8776
40	0.2033	0.9293	0.4453	0.8752	0.875
50	0.1901	0.9336	0.4928	0.878	0.8779
60	0.1659	0.9379	0.517	0.8763	0.876
70	0.1518	0.9467	0.5608	0.8803	0.8801
80	0.1367	0.9476	0.6003	0.8789	0.8793
90	0.1276	0.9491	0.652	0.8751	0.8742
100	0.1159	0.9576	0.7036	0.8788	0.8783
110	0.109	0.9594	0.7391	0.8772	0.8764
120	0.1006	0.9631	0.7741	0.8781	0.8782
130	0.0953	0.9651	0.8203	0.8787	0.8785
140	0.0906	0.9672	0.862	0.8767	0.8766
150	0.0862	0.9683	0.8971	0.8768	0.8766
160	0.0829	0.9689	0.9071	0.8762	0.8762
170	0.0802	0.9706	0.9382	0.8758	0.8758
180	0.0781	0.9707	0.957	0.8763	0.8764
190	0.0761	0.9718	0.9886	0.8772	0.8771
200	0.0748	0.9721	1.0011	0.8761	0.8761
210	0.0737	0.9727	1.0196	0.8768	0.8768
220	0.0724	0.9728	1.038	0.8764	0.8763
230	0.0717	0.9732	1.0463	0.8764	0.8763
240	0.071	0.9736	1.0541	0.8759	0.8758
250	0.0705	0.9737	1.064	0.8757	0.8756
260	0.07	0.9736	1.0695	0.8761	0.876
270	0.0697	0.9737	1.0741	0.8751	0.875
280	0.0694	0.9738	1.0793	0.8752	0.8751
290	0.0691	0.9739	1.0835	0.8751	0.8749
300	0.0689	0.974	1.087	0.8752	0.875
310	0.0688	0.9739	1.0893	0.875	0.8748
320	0.0686	0.9741	1.0911	0.875	0.8749
330	0.0685	0.974	1.0918	0.8746	0.8744
340	0.0684	0.9741	1.0944	0.8749	0.8747
350	0.0683	0.974	1.0954	0.8745	0.8744
360	0.0683	0.9741	1.0967	0.8747	0.8745
370	0.0682	0.974	1.0968	0.8744	0.8743
380	0.0682	0.9741	1.0976	0.8747	0.8745
390	0.0681	0.9741	1.0982	0.8747	0.8745
400	0.0681	0.974	1.0986	0.8746	0.8744
410	0.0681	0.9741	1.0989	0.8744	0.8743
420	0.0681	0.9741	1.0992	0.8747	0.8745
430	0.068	0.9741	1.0994	0.8746	0.8744
440	0.068	0.9741	1.0995	0.8746	0.8744
450	0.068	0.9741	1.0996	0.8745	0.8744
460	0.068	0.9741	1.0998	0.8745	0.8744
470	0.068	0.9741	1.0999	0.8745	0.8744
480	0.068	0.9741	1.1	0.8745	0.8744
490	0.068	0.9741	1.1	0.8745	0.8744
500	0.068	0.9741	1.1001	0.8745	0.8744

Plots

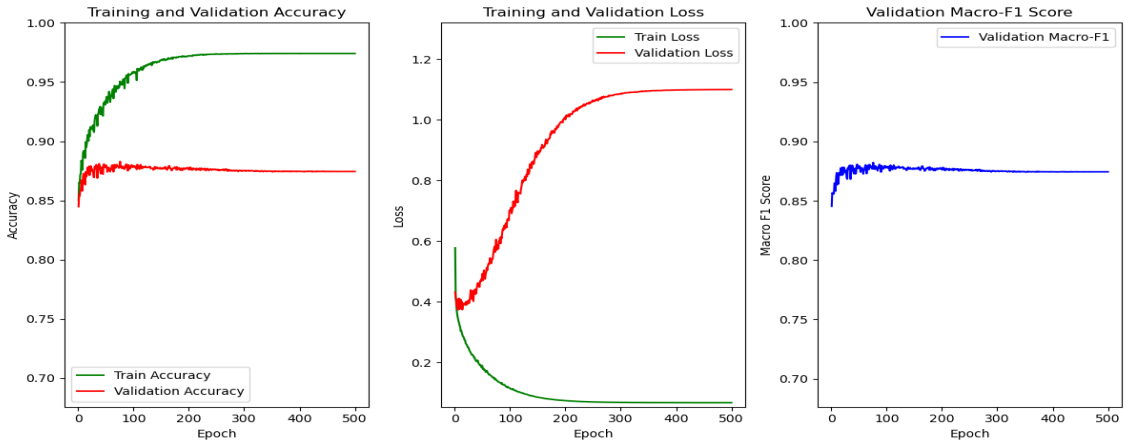

Figure 5: Various Plots

Confusion Matrix on Validation Data

Actual \ Predicted	Class 0	Class 1	Class 2	Class 3	Class 4	Class 5	Class 6	Class 7	Class 8	Class 9
Class 0	1022	5	22	45	5	0	101	0	2	0
Class 1	0	1195	1	15	5	1	2	0	0	0
Class 2	12	3	990	13	112	0	69	0	6	0
Class 3	31	19	19	1055	35	0	20	0	5	0
Class 4	2	2	96	36	1010	1	48	0	7	0
Class 5	1	0	0	1	0	1157	1	33	5	13
Class 6	189	0	104	43	94	1	776	0	11	0
Class 7	0	0	0	0	0	14	0	1102	3	40
Class 8	3	1	9	5	10	4	18	3	1143	1
Class 9	0	0	0	0	0	12	0	45	1	1145

Table 4: Confusion Matrix

2.2 Model 1 (A Bit Deep Model without Dropout)

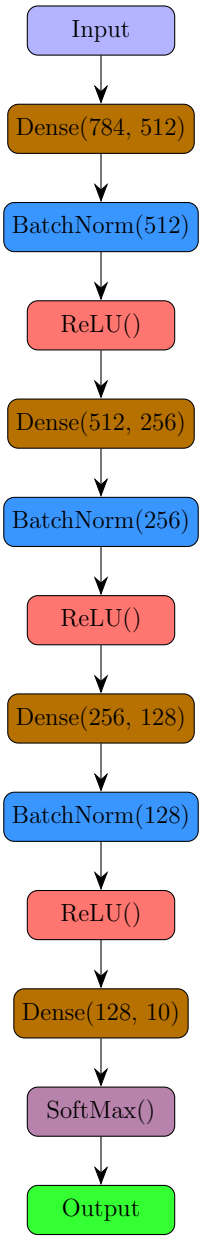

Figure 6: **Model 1 Architecture**

2.2.1 Learning Rate 0.0005

Performance Metrics Table

Epoch	Train Loss	Train Accuracy	Validation Loss	Validation Accuracy	Validation Macro-F1
10	0.145	0.9557	0.3129	0.8948	0.8952
20	0.0622	0.9875	0.3819	0.8968	0.8974
30	0.0308	0.9876	0.484	0.8944	0.8938
40	0.0146	0.9956	0.5181	0.8992	0.8988
50	0.0087	0.9985	0.5719	0.8934	0.8935
60	0.0054	0.9979	0.5885	0.9008	0.9011
70	0.0047	1.0	0.6197	0.901	0.901
80	0.0025	0.9996	0.6269	0.9013	0.9014
90	0.0026	1.0	0.6527	0.9022	0.9021
100	0.0013	1.0	0.6797	0.9021	0.9023
110	0.0007	1.0	0.6886	0.9012	0.9012
120	0.0004	0.9998	0.6888	0.9032	0.9034
130	0.0002	1.0	0.7115	0.9044	0.9044
140	0.0003	1.0	0.7191	0.9042	0.9042
150	0.0003	1.0	0.721	0.9048	0.9047
160	0.0002	1.0	0.7192	0.9045	0.9045
170	0.0001	1.0	0.7332	0.9042	0.9041
180	0.0001	1.0	0.7283	0.9052	0.9053
190	0.0001	1.0	0.7359	0.9046	0.9046
200	0.0001	1.0	0.7459	0.9048	0.9047
210	0.0001	1.0	0.7374	0.9059	0.9059
220	0.0001	1.0	0.7408	0.9047	0.9046
230	0.0	1.0	0.7486	0.9043	0.9043
240	0.0	1.0	0.7502	0.9067	0.9068
250	0.0	1.0	0.7451	0.9058	0.9057
260	0.0	1.0	0.7509	0.9053	0.9055
270	0.0	1.0	0.7432	0.9056	0.9056
280	0.0	1.0	0.7448	0.9055	0.9056
290	0.0	1.0	0.749	0.9059	0.906
300	0.0	1.0	0.7519	0.9052	0.9055
310	0.0	1.0	0.7542	0.9061	0.9061
320	0.0	1.0	0.751	0.9063	0.9062
330	0.0	1.0	0.7529	0.905	0.905
340	0.0	1.0	0.7532	0.9058	0.9057
350	0.0	1.0	0.7612	0.9049	0.905
360	0.0	1.0	0.7522	0.9062	0.9063
370	0.0	1.0	0.7488	0.9061	0.9062
380	0.0	1.0	0.7559	0.9054	0.9055
390	0.0	1.0	0.7563	0.9056	0.9056
400	0.0	1.0	0.7523	0.9058	0.9058
410	0.0	1.0	0.7594	0.9052	0.9053
420	0.0	1.0	0.7538	0.9052	0.9053
430	0.0	1.0	0.7528	0.9056	0.9056
440	0.0	1.0	0.7497	0.906	0.906
450	0.0	1.0	0.7528	0.9058	0.9059
460	0.0	1.0	0.7532	0.9053	0.9054
470	0.0	1.0	0.7553	0.9053	0.9052
480	0.0	1.0	0.7543	0.9053	0.9054
490	0.0	1.0	0.7545	0.9065	0.9065
500	0.0	1.0	0.7528	0.9048	0.9048

Plots

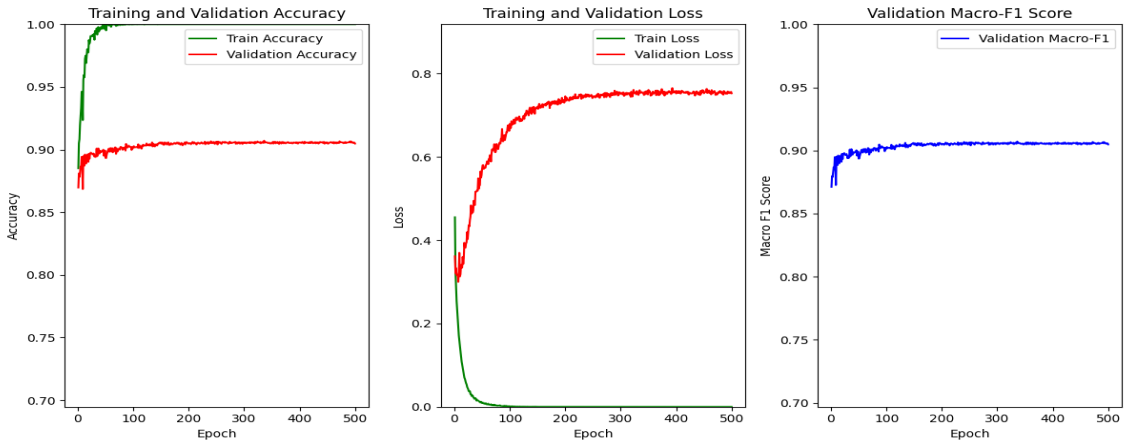

Figure 7: Various Plots

Confusion Matrix on Validation Data

Actual \ Predicted	Class 0	Class 1	Class 2	Class 3	Class 4	Class 5	Class 6	Class 7	Class 8	Class 9
Class 0	1044	1	17	25	7	2	102	0	4	0
Class 1	2	1198	2	12	1	0	3	0	1	0
Class 2	21	0	1044	5	74	0	60	0	1	0
Class 3	26	6	12	1078	31	0	25	0	4	2
Class 4	7	2	81	36	1018	0	49	1	7	1
Class 5	0	0	0	1	0	1174	0	17	2	17
Class 6	135	2	84	34	50	1	904	0	8	0
Class 7	0	0	0	0	0	14	0	1109	2	34
Class 8	8	0	11	2	7	3	11	3	1150	2
Class 9	0	0	0	1	0	11	0	32	1	1158

Table 5: Confusion Matrix

2.2.2 Learning Rate 0.001

Performance Metrics Table

Epoch	Train Loss	Train Accuracy	Validation Loss	Validation Accuracy	Validation Macro-F1
10	0.1638	0.9479	0.318	0.8951	0.895
20	0.0735	0.9866	0.3734	0.9002	0.9006
30	0.0351	0.9843	0.5119	0.8932	0.8928
40	0.0209	0.9924	0.5624	0.8942	0.8937
50	0.0105	0.9991	0.5717	0.9018	0.902
60	0.008	0.9989	0.6143	0.8984	0.8984
70	0.0058	0.9986	0.6661	0.8965	0.8971
80	0.0037	0.9999	0.6669	0.9002	0.9003
90	0.0018	0.9998	0.7029	0.9002	0.9003
100	0.0016	1.0	0.71	0.9032	0.9033
110	0.0012	1.0	0.7313	0.9028	0.9027
120	0.0006	1.0	0.7365	0.9021	0.902
130	0.0006	1.0	0.7513	0.9013	0.9012
140	0.0004	1.0	0.7528	0.902	0.9018
150	0.0003	1.0	0.7787	0.9022	0.902
160	0.0002	1.0	0.767	0.9028	0.9027
170	0.0001	1.0	0.7728	0.9037	0.9035
180	0.0001	1.0	0.7772	0.9035	0.9033
190	0.0001	1.0	0.7909	0.9023	0.9021
200	0.0001	1.0	0.7915	0.9042	0.9041
210	0.0001	1.0	0.7897	0.9026	0.9025
220	0.0001	1.0	0.7973	0.9032	0.903
230	0.0	1.0	0.7984	0.903	0.9029
240	0.0	1.0	0.7957	0.9039	0.9038
250	0.0	1.0	0.7998	0.9033	0.9031
260	0.0	1.0	0.802	0.9046	0.9045
270	0.0	1.0	0.7988	0.9041	0.9039
280	0.0	1.0	0.7987	0.9037	0.9036
290	0.0	1.0	0.8039	0.9045	0.9044
300	0.0	1.0	0.8002	0.9037	0.9037
310	0.0	1.0	0.8037	0.9038	0.9036
320	0.0	1.0	0.8052	0.9038	0.9037
330	0.0	1.0	0.8082	0.9037	0.9035
340	0.0	1.0	0.8066	0.9048	0.9046
350	0.0001	1.0	0.8132	0.9029	0.9027
360	0.0	1.0	0.8065	0.9036	0.9034
370	0.0	1.0	0.799	0.9051	0.9051
380	0.0	1.0	0.8084	0.9032	0.9032
390	0.0	1.0	0.807	0.9031	0.9028
400	0.0	1.0	0.8022	0.9038	0.9037
410	0.0	1.0	0.8133	0.903	0.9029
420	0.0	1.0	0.8062	0.9041	0.904
430	0.0	1.0	0.8038	0.9036	0.9035
440	0.0	1.0	0.8007	0.9042	0.904
450	0.0	1.0	0.8001	0.9035	0.9034
460	0.0	1.0	0.8066	0.9046	0.9044
470	0.0	1.0	0.8072	0.904	0.9037
480	0.0	1.0	0.8079	0.904	0.9038
490	0.0	1.0	0.807	0.9039	0.9038
500	0.0	1.0	0.8032	0.9029	0.9027

Plots

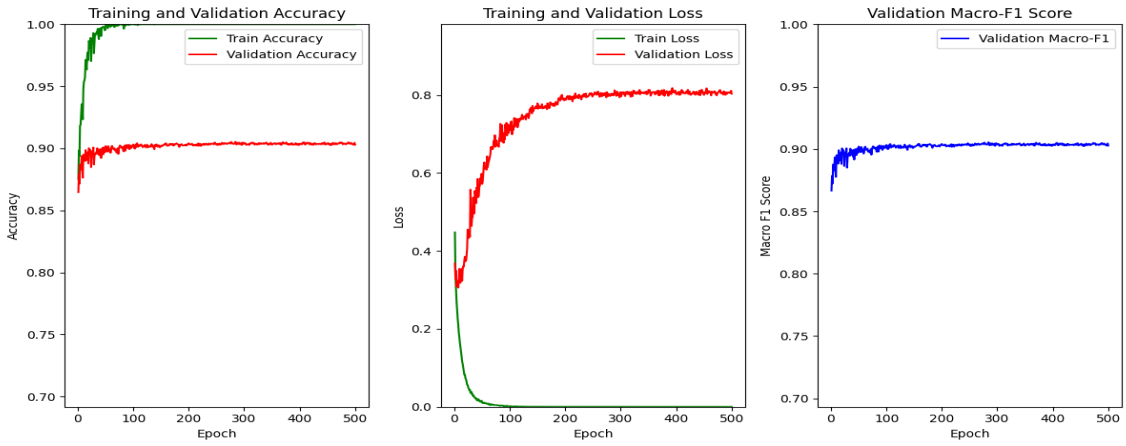

Figure 8: Various Plots

Confusion Matrix on Validation Data

Actual \ Predicted	Class 0	Class 1	Class 2	Class 3	Class 4	Class 5	Class 6	Class 7	Class 8	Class 9
Class 0	1042	1	20	25	5	1	102	0	6	0
Class 1	1	1197	1	13	4	0	2	0	1	0
Class 2	21	0	1019	6	89	1	67	0	2	0
Class 3	27	10	7	1072	31	0	27	0	7	3
Class 4	8	1	67	35	1028	0	56	1	6	0
Class 5	0	0	0	0	0	1177	0	16	2	16
Class 6	137	3	84	25	55	0	900	0	13	1
Class 7	0	0	0	0	0	13	0	1114	4	28
Class 8	6	2	7	2	8	3	9	3	1155	2
Class 9	0	0	0	0	0	11	0	33	1	1158

Table 6: Confusion Matrix

2.2.3 Learning Rate 0.005

Performance Metrics Table

Epoch	Train Loss	Train Accuracy	Validation Loss	Validation Accuracy	Validation Macro-F1
10	0.2001	0.9413	0.3022	0.8962	0.8963
20	0.1031	0.9741	0.378	0.8969	0.8976
30	0.0559	0.9854	0.4897	0.8991	0.899
40	0.0334	0.9908	0.577	0.8949	0.8947
50	0.0218	0.9962	0.6391	0.8982	0.898
60	0.015	0.9988	0.6675	0.9006	0.9007
70	0.01	0.9988	0.7234	0.9025	0.9025
80	0.0072	0.9998	0.7603	0.9032	0.9034
90	0.0044	1.0	0.7866	0.9024	0.9026
100	0.0037	0.9995	0.8338	0.901	0.9012
110	0.002	0.9994	0.856	0.9011	0.9011
120	0.0012	1.0	0.8481	0.9043	0.9043
130	0.0016	1.0	0.8764	0.9049	0.9048
140	0.0007	1.0	0.8932	0.9036	0.9035
150	0.0006	1.0	0.8936	0.9043	0.9044
160	0.0003	1.0	0.8926	0.9062	0.9064
170	0.0004	1.0	0.9198	0.9038	0.9038
180	0.0003	1.0	0.8991	0.9058	0.9058
190	0.0002	1.0	0.9233	0.9042	0.9043
200	0.0002	1.0	0.9232	0.9049	0.9049
210	0.0003	1.0	0.9371	0.9038	0.9037
220	0.0001	1.0	0.9372	0.9052	0.9052
230	0.0001	1.0	0.9317	0.9066	0.9066
240	0.0002	1.0	0.9279	0.9056	0.9056
250	0.0002	1.0	0.9304	0.9065	0.9064
260	0.0001	1.0	0.9331	0.9071	0.9073
270	0.0002	1.0	0.9359	0.9062	0.9061
280	0.0001	1.0	0.928	0.9075	0.9076
290	0.0002	1.0	0.9351	0.9066	0.9066
300	0.0001	1.0	0.9376	0.9065	0.9067
310	0.0001	1.0	0.9415	0.9066	0.9065
320	0.0001	1.0	0.9374	0.9058	0.9056
330	0.0001	1.0	0.9347	0.9063	0.9063
340	0.0001	1.0	0.933	0.9056	0.9055
350	0.0001	1.0	0.9409	0.9057	0.9056
360	0.0001	1.0	0.9347	0.9069	0.9069
370	0.0001	1.0	0.933	0.9058	0.9059
380	0.0	1.0	0.9396	0.9064	0.9064
390	0.0001	1.0	0.9414	0.906	0.9059
400	0.0001	1.0	0.9368	0.9057	0.9057
410	0.0001	1.0	0.9458	0.906	0.9059
420	0.0001	1.0	0.9396	0.9063	0.9064
430	0.0001	1.0	0.933	0.9062	0.906
440	0.0002	1.0	0.9266	0.9061	0.9061
450	0.0	1.0	0.9371	0.9067	0.9068
460	0.0	1.0	0.9358	0.9068	0.9069
470	0.0001	1.0	0.936	0.9068	0.9067
480	0.0002	1.0	0.9382	0.9066	0.9066
490	0.0001	1.0	0.9423	0.9063	0.9062
500	0.0	1.0	0.9393	0.9061	0.906

Plots

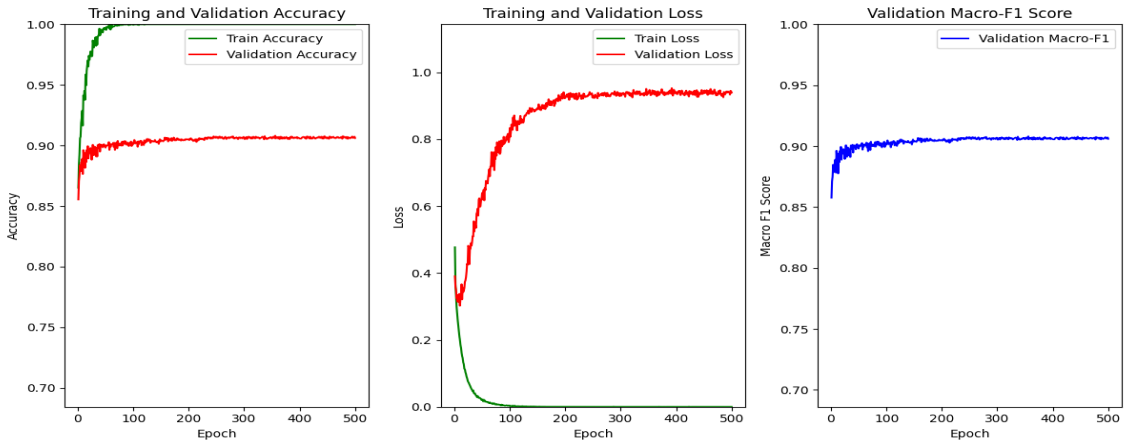

Figure 9: Various Plots

Confusion Matrix on Validation Data

Actual \ Predicted	Class 0	Class 1	Class 2	Class 3	Class 4	Class 5	Class 6	Class 7	Class 8	Class 9
Class 0	1055	0	19	37	3	0	84	0	4	0
Class 1	2	1195	1	14	2	0	4	0	1	0
Class 2	24	0	1038	8	68	0	66	0	1	0
Class 3	30	7	12	1079	29	0	18	0	6	3
Class 4	5	1	72	37	1032	1	48	1	5	0
Class 5	0	0	0	0	0	1178	0	15	2	16
Class 6	146	0	92	28	53	0	893	0	5	1
Class 7	0	0	0	0	0	13	0	1109	2	35
Class 8	6	0	5	1	6	2	14	6	1156	1
Class 9	0	0	0	0	0	11	0	38	2	1152

Table 7: Confusion Matrix

2.2.4 Learning Rate 0.01

Performance Metrics Table

Epoch	Train Loss	Train Accuracy	Validation Loss	Validation Accuracy	Validation Macro-F1
10	0.2144	0.9323	0.3122	0.8911	0.8919
20	0.1192	0.9658	0.3859	0.8934	0.894
30	0.0686	0.9754	0.509	0.8932	0.893
40	0.0413	0.9881	0.5806	0.8953	0.8946
50	0.0257	0.9957	0.6563	0.897	0.8968
60	0.0184	0.9977	0.7177	0.8953	0.8949
70	0.0125	0.9996	0.7425	0.8988	0.8991
80	0.009	0.9994	0.796	0.8963	0.8965
90	0.0062	0.9976	0.8375	0.8979	0.8978
100	0.0058	0.9995	0.8678	0.8982	0.899
110	0.003	0.9999	0.9205	0.8991	0.899
120	0.0017	1.0	0.922	0.8994	0.8997
130	0.0012	1.0	0.9276	0.9008	0.9009
140	0.0011	1.0	0.9471	0.9018	0.9018
150	0.0005	1.0	0.9877	0.8992	0.8989
160	0.0003	1.0	0.9772	0.9002	0.9006
170	0.0007	1.0	0.9914	0.9013	0.9014
180	0.0006	1.0	0.9849	0.9013	0.9014
190	0.0005	1.0	0.9983	0.9005	0.9005
200	0.0002	1.0	0.9939	0.9018	0.9018
210	0.0003	1.0	1.0125	0.9018	0.9017
220	0.0003	1.0	1.0128	0.9012	0.9013
230	0.0001	1.0	1.0103	0.9015	0.9016
240	0.0002	1.0	1.013	0.9018	0.9019
250	0.0002	1.0	1.0126	0.9026	0.9025
260	0.0004	1.0	1.0134	0.9031	0.9032
270	0.0003	1.0	1.0028	0.9034	0.9033
280	0.0001	1.0	1.0064	0.9016	0.9018
290	0.0001	1.0	1.0178	0.902	0.9021
300	0.0001	1.0	1.0292	0.9021	0.9023
310	0.0001	1.0	1.0223	0.9032	0.9032
320	0.0001	1.0	1.0245	0.902	0.902
330	0.0001	1.0	1.0193	0.9026	0.9027
340	0.0002	1.0	1.0206	0.9024	0.9025
350	0.0004	1.0	1.0341	0.9022	0.9022
360	0.0002	1.0	1.0219	0.9013	0.9014
370	0.0002	1.0	1.027	0.9025	0.9026
380	0.0	1.0	1.0312	0.9023	0.9024
390	0.0001	1.0	1.0278	0.9016	0.9016
400	0.0001	1.0	1.0283	0.9022	0.9024
410	0.0002	1.0	1.0306	0.9023	0.9024
420	0.0001	1.0	1.0201	0.9023	0.9025
430	0.0001	1.0	1.0235	0.9023	0.9024
440	0.0001	1.0	1.0179	0.9033	0.9034
450	0.0001	1.0	1.0199	0.9025	0.9026
460	0.0001	1.0	1.0313	0.9028	0.9028
470	0.0001	1.0	1.0188	0.9034	0.9034
480	0.0001	1.0	1.0274	0.9022	0.9022
490	0.0001	1.0	1.0327	0.9023	0.9024
500	0.0001	1.0	1.0212	0.9027	0.9027

Plots

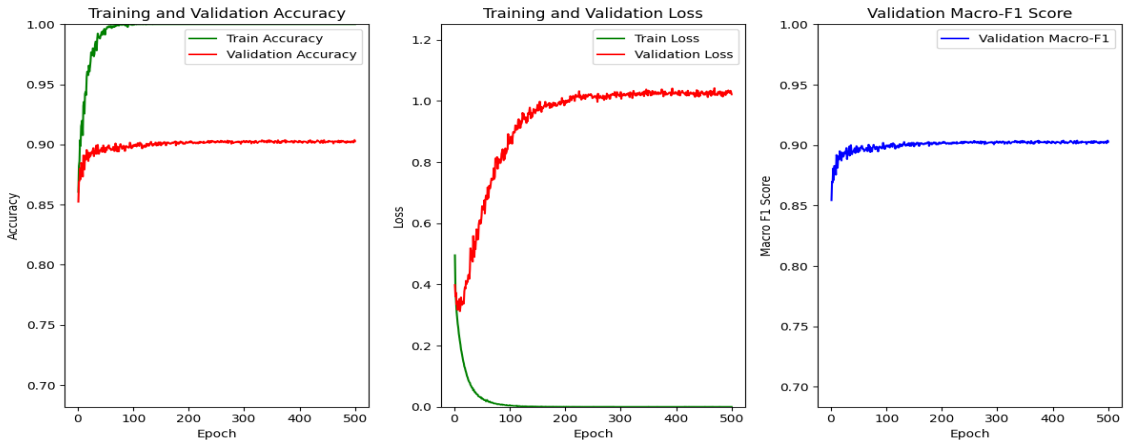

Figure 10: Various Plots

Confusion Matrix on Validation Data

Actual \ Predicted	Class 0	Class 1	Class 2	Class 3	Class 4	Class 5	Class 6	Class 7	Class 8	Class 9
Class 0	1041	1	20	29	5	0	103	0	3	0
Class 1	0	1192	2	17	2	0	4	0	2	0
Class 2	20	1	1034	8	83	0	55	0	4	0
Class 3	25	8	9	1086	24	2	25	0	5	0
Class 4	5	1	94	36	1011	0	50	0	4	1
Class 5	0	0	0	0	0	1176	0	17	2	16
Class 6	150	0	99	31	49	0	880	0	8	1
Class 7	0	0	0	0	0	20	0	1112	1	26
Class 8	4	3	6	6	6	5	12	2	1152	1
Class 9	0	0	0	0	0	11	0	32	1	1159

Table 8: Confusion Matrix

2.3 Model 2 (Deep Model with Every Layer)

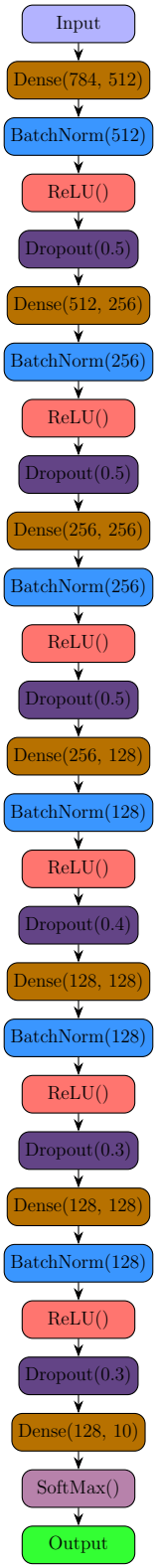

Figure 11: **Model 2 Architecture**

2.3.1 Learning Rate 0.0005

Performance Metrics Table

Epoch	Train Loss	Train Accuracy	Validation Loss	Validation Accuracy	Validation Macro-F1
10	0.4483	0.8888	0.3632	0.8734	0.8733
20	0.3699	0.9112	0.3262	0.8873	0.8876
30	0.3213	0.9239	0.3107	0.8905	0.8911
40	0.2878	0.9352	0.2991	0.8964	0.8969
50	0.2668	0.9415	0.2987	0.8983	0.8984
60	0.247	0.9513	0.2947	0.9002	0.9004
70	0.2282	0.9559	0.2945	0.9018	0.9022
80	0.2124	0.961	0.2979	0.9038	0.9036
90	0.2025	0.965	0.3062	0.9035	0.9035
100	0.1915	0.9693	0.2974	0.9042	0.9043
110	0.1828	0.971	0.3016	0.9031	0.9031
120	0.1802	0.9722	0.3059	0.9045	0.9048
130	0.1725	0.9744	0.3098	0.9048	0.9047
140	0.1689	0.976	0.3101	0.9062	0.9064
150	0.1635	0.9769	0.3123	0.9045	0.9043
160	0.164	0.9776	0.3145	0.9058	0.9058
170	0.1595	0.9788	0.313	0.9056	0.9057
180	0.1568	0.9788	0.3152	0.9046	0.9043
190	0.1569	0.9791	0.3201	0.9039	0.9043
200	0.1568	0.9792	0.325	0.9051	0.9054
210	0.155	0.9801	0.3198	0.9061	0.9062
220	0.1516	0.9804	0.3206	0.9062	0.9061
230	0.151	0.9803	0.322	0.9057	0.9059
240	0.1492	0.981	0.3169	0.9058	0.9058
250	0.1525	0.9808	0.3222	0.905	0.9051
260	0.1484	0.9808	0.3211	0.9049	0.905
270	0.1472	0.9813	0.3207	0.9054	0.9057
280	0.1491	0.9809	0.3237	0.905	0.9051
290	0.1516	0.9808	0.3211	0.9048	0.9043
300	0.1496	0.9816	0.3218	0.906	0.9059
310	0.1502	0.9811	0.3215	0.9058	0.906
320	0.1507	0.9811	0.3208	0.9046	0.9046
330	0.1508	0.9811	0.3225	0.9058	0.9059
340	0.1466	0.9818	0.3217	0.9052	0.9051
350	0.1504	0.9815	0.32	0.906	0.9061
360	0.1465	0.9816	0.3199	0.9052	0.9054
370	0.1508	0.9814	0.319	0.9053	0.9054
380	0.149	0.9812	0.319	0.9055	0.9053
390	0.1482	0.9812	0.3238	0.9043	0.9047
400	0.1466	0.9821	0.3219	0.9046	0.9047
410	0.1493	0.9813	0.3231	0.9042	0.904
420	0.15	0.9819	0.3293	0.9039	0.9041
430	0.1486	0.9812	0.324	0.906	0.9059
440	0.1444	0.9811	0.3194	0.9062	0.9061
450	0.147	0.9816	0.3198	0.9048	0.9051
460	0.1522	0.9813	0.3217	0.9054	0.9054
470	0.1537	0.9808	0.3221	0.906	0.9059
480	0.1453	0.9813	0.3188	0.9049	0.9051
490	0.148	0.9815	0.3184	0.9053	0.9057
500	0.149	0.9821	0.3243	0.9048	0.9051

Plots

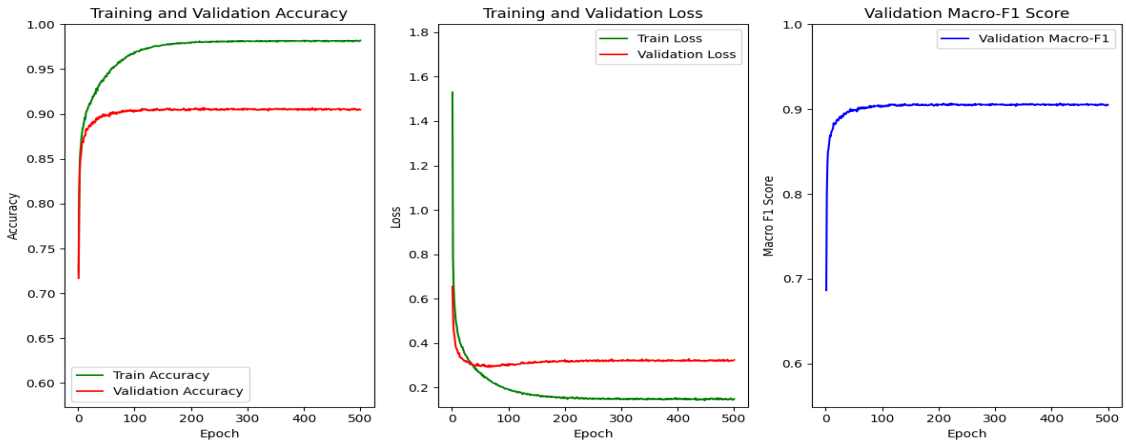

Figure 12: Various Plots

Confusion Matrix on Validation Data

Actual \ Predicted	Class 0	Class 1	Class 2	Class 3	Class 4	Class 5	Class 6	Class 7	Class 8	Class 9
Class 0	1026	0	22	37	2	0	113	0	2	0
Class 1	0	1198	2	15	2	0	1	0	1	0
Class 2	21	0	1048	7	66	0	62	0	1	0
Class 3	25	6	3	1101	28	0	19	0	2	0
Class 4	1	1	93	37	1000	0	66	0	4	0
Class 5	0	0	0	0	0	1182	0	18	0	11
Class 6	140	0	86	34	51	0	900	1	6	0
Class 7	0	0	0	0	0	18	0	1112	2	27
Class 8	2	2	4	4	5	1	16	4	1158	1
Class 9	0	0	0	0	0	11	0	37	1	1154

Table 9: Confusion Matrix

2.3.2 Learning Rate 0.001

Performance Metrics Table

Epoch	Train Loss	Train Accuracy	Validation Loss	Validation Accuracy	Validation Macro-F1
10	0.4235	0.893	0.3508	0.8751	0.8749
20	0.349	0.9168	0.3121	0.8909	0.8904
30	0.2999	0.9263	0.3027	0.8944	0.8951
40	0.2679	0.9404	0.2923	0.8997	0.9
50	0.2412	0.9491	0.2906	0.8998	0.8993
60	0.2229	0.9581	0.294	0.9018	0.9013
70	0.2013	0.9624	0.2947	0.9033	0.9033
80	0.1864	0.9681	0.2979	0.9038	0.9034
90	0.1771	0.9721	0.3015	0.903	0.9026
100	0.1672	0.9752	0.3016	0.9046	0.9046
110	0.1524	0.9774	0.3126	0.9045	0.9041
120	0.1508	0.9795	0.3149	0.9054	0.9054
130	0.1456	0.9809	0.314	0.9047	0.9046
140	0.1379	0.982	0.3197	0.9051	0.9049
150	0.1356	0.9838	0.3218	0.9056	0.9053
160	0.1344	0.9841	0.3254	0.9049	0.9048
170	0.1307	0.9854	0.3245	0.9055	0.9054
180	0.1271	0.9852	0.3222	0.9059	0.9058
190	0.125	0.986	0.3307	0.9048	0.9046
200	0.1265	0.9858	0.3326	0.9038	0.9041
210	0.1253	0.9864	0.3223	0.9061	0.9061
220	0.1227	0.9868	0.3301	0.9061	0.906
230	0.1238	0.9866	0.3271	0.9061	0.9061
240	0.1216	0.9871	0.3288	0.9052	0.9048
250	0.1213	0.9871	0.3282	0.9045	0.9045
260	0.119	0.9871	0.3276	0.9049	0.9048
270	0.1218	0.9871	0.33	0.9066	0.9065
280	0.1193	0.987	0.3344	0.9056	0.9053
290	0.1214	0.9871	0.3311	0.905	0.9047
300	0.1196	0.9872	0.3313	0.9065	0.9062
310	0.1213	0.9871	0.3334	0.9045	0.904
320	0.1213	0.9874	0.3298	0.9065	0.9063
330	0.1222	0.9871	0.3307	0.9059	0.9058
340	0.119	0.9874	0.3307	0.9067	0.9066
350	0.1212	0.9872	0.3293	0.9052	0.9053
360	0.1214	0.9873	0.3238	0.9057	0.9058
370	0.1221	0.9871	0.3304	0.9056	0.9053
380	0.1174	0.9875	0.3286	0.9056	0.9054
390	0.1185	0.9874	0.332	0.9052	0.9052
400	0.1172	0.9874	0.3385	0.9055	0.905
410	0.1215	0.9874	0.3316	0.9057	0.9052
420	0.1238	0.9873	0.3345	0.9058	0.9058
430	0.1205	0.9869	0.3332	0.9058	0.9054
440	0.1156	0.9871	0.3314	0.9061	0.9059
450	0.119	0.9873	0.328	0.9062	0.9062
460	0.1181	0.9874	0.3303	0.9058	0.9059
470	0.1213	0.9871	0.3367	0.9054	0.9049
480	0.1169	0.9877	0.3295	0.9042	0.9042
490	0.1205	0.9875	0.328	0.906	0.9059
500	0.1201	0.9875	0.3367	0.9053	0.9053

Plots

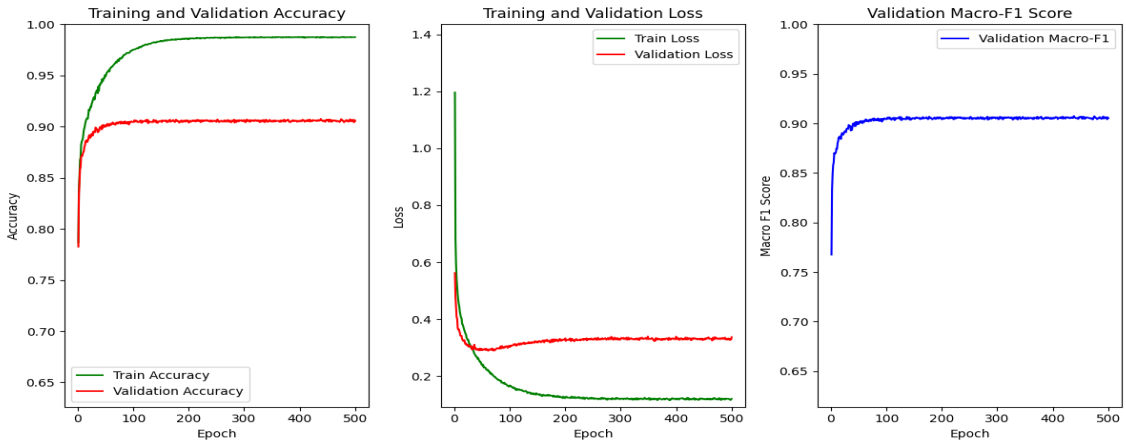

Figure 13: Various Plots

Confusion Matrix on Validation Data

Actual \ Predicted	Class 0	Class 1	Class 2	Class 3	Class 4	Class 5	Class 6	Class 7	Class 8	Class 9
Class 0	1033	2	23	37	2	0	103	0	2	0
Class 1	2	1193	4	16	1	0	2	0	1	0
Class 2	21	0	1060	8	65	0	50	0	1	0
Class 3	27	7	6	1100	30	0	12	0	2	0
Class 4	0	1	89	37	1010	0	61	0	4	0
Class 5	0	0	0	0	0	1187	0	14	0	10
Class 6	140	1	99	34	50	0	887	0	7	0
Class 7	0	0	0	0	0	20	0	1107	2	30
Class 8	3	1	4	4	5	2	16	3	1157	2
Class 9	0	0	0	0	0	15	0	36	1	1151

Table 10: Confusion Matrix

2.3.3 Learning Rate 0.005

Performance Metrics Table

Epoch	Train Loss	Train Accuracy	Validation Loss	Validation Accuracy	Validation Macro-F1
10	0.4198	0.8976	0.34	0.881	0.8809
20	0.3436	0.9179	0.3044	0.8934	0.8935
30	0.2936	0.9323	0.2962	0.897	0.8976
40	0.2616	0.9437	0.2872	0.9012	0.9015
50	0.2342	0.9526	0.2873	0.9018	0.9014
60	0.2147	0.9593	0.2832	0.9038	0.903
70	0.1958	0.9648	0.2862	0.9037	0.9033
80	0.1795	0.972	0.2881	0.9038	0.9034
90	0.1691	0.9747	0.2971	0.9052	0.9049
100	0.1599	0.9785	0.2911	0.9058	0.9059
110	0.1461	0.9803	0.3038	0.9077	0.9075
120	0.1412	0.9824	0.3075	0.9051	0.9053
130	0.1365	0.9839	0.3038	0.9072	0.9073
140	0.1305	0.985	0.3069	0.9063	0.9064
150	0.127	0.9858	0.313	0.907	0.9068
160	0.1243	0.9869	0.3138	0.9066	0.9064
170	0.1238	0.9876	0.3133	0.9074	0.9074
180	0.1193	0.9881	0.3157	0.9065	0.9066
190	0.1189	0.9879	0.3149	0.9062	0.9063
200	0.1142	0.9878	0.3228	0.9068	0.9073
210	0.1186	0.9888	0.3097	0.907	0.9069
220	0.1136	0.9887	0.3217	0.9072	0.9073
230	0.1146	0.9889	0.3232	0.9062	0.9064
240	0.113	0.9891	0.3201	0.9067	0.9065
250	0.1126	0.9892	0.3171	0.9062	0.9063
260	0.1111	0.9892	0.3216	0.9058	0.9058
270	0.1111	0.9895	0.3201	0.9061	0.9061
280	0.1096	0.9894	0.3253	0.9059	0.9061
290	0.1118	0.9895	0.3175	0.9064	0.9064
300	0.1136	0.9897	0.3265	0.9063	0.9063
310	0.1121	0.9894	0.3218	0.9061	0.906
320	0.1106	0.9896	0.3232	0.9068	0.9069
330	0.1109	0.9893	0.3265	0.9055	0.9056
340	0.1127	0.9899	0.3245	0.9061	0.9061
350	0.1116	0.9896	0.3224	0.9074	0.9074
360	0.1104	0.9897	0.3184	0.9065	0.9067
370	0.1105	0.9896	0.3225	0.907	0.9068
380	0.1114	0.9896	0.319	0.9067	0.9068
390	0.1102	0.9895	0.3257	0.9067	0.9066
400	0.1091	0.9898	0.3255	0.9072	0.907
410	0.1118	0.9898	0.3238	0.9055	0.9053
420	0.1106	0.9896	0.3269	0.9063	0.9062
430	0.107	0.9893	0.3347	0.9052	0.9048
440	0.1076	0.9895	0.3207	0.9056	0.9055
450	0.1095	0.9895	0.3192	0.9073	0.9075
460	0.1109	0.9896	0.3301	0.9054	0.9052
470	0.1117	0.9893	0.321	0.9059	0.906
480	0.1066	0.9899	0.3213	0.9065	0.9066
490	0.1116	0.9898	0.3225	0.9064	0.9067
500	0.1103	0.9895	0.3291	0.9063	0.9062

Plots

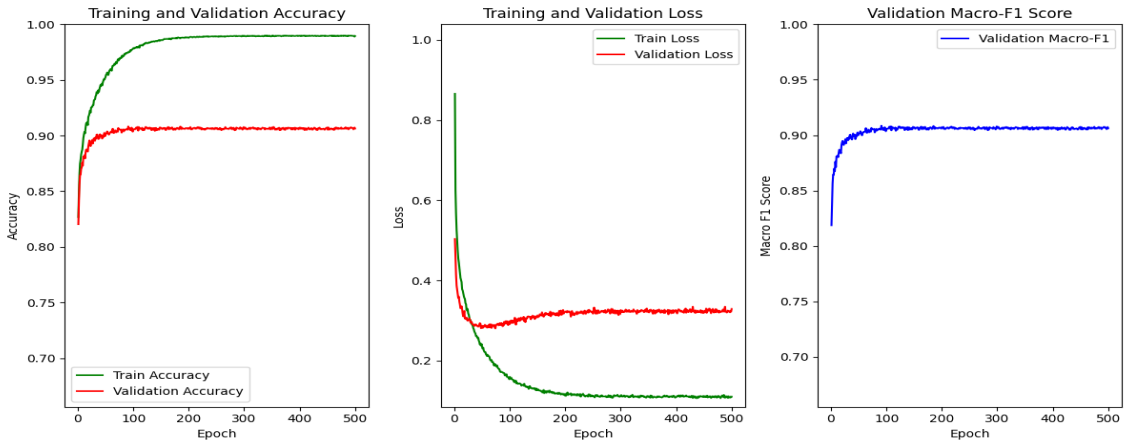

Figure 14: Various Plots

Confusion Matrix on Validation Data

Actual \ Predicted	Class 0	Class 1	Class 2	Class 3	Class 4	Class 5	Class 6	Class 7	Class 8	Class 9
Class 0	1029	1	22	29	3	0	116	0	2	0
Class 1	1	1195	2	16	2	0	2	0	1	0
Class 2	14	0	1061	5	62	0	61	0	2	0
Class 3	27	6	12	1078	34	0	25	0	2	0
Class 4	0	1	98	28	1006	0	65	0	4	0
Class 5	0	0	0	0	0	1178	0	19	1	13
Class 6	136	0	93	21	50	0	912	0	6	0
Class 7	0	0	0	0	0	17	0	1117	0	25
Class 8	3	1	4	2	4	2	16	2	1161	2
Class 9	0	0	0	0	0	11	0	38	1	1153

Table 11: Confusion Matrix

2.3.4 Learning Rate 0.01

Performance Metrics Table

Epoch	Train Loss	Train Accuracy	Validation Loss	Validation Accuracy	Validation Macro-F1
10	0.441	0.8908	0.3593	0.874	0.8715
20	0.3623	0.9137	0.3111	0.8907	0.8914
30	0.308	0.9275	0.2946	0.8989	0.8988
40	0.2749	0.941	0.2928	0.8985	0.8986
50	0.2474	0.9495	0.2875	0.9022	0.9026
60	0.2256	0.9557	0.2827	0.9038	0.9034
70	0.2069	0.9622	0.2864	0.904	0.9039
80	0.1835	0.9682	0.2891	0.9046	0.9043
90	0.1773	0.9707	0.2878	0.9039	0.9031
100	0.1695	0.9748	0.2896	0.906	0.9056
110	0.1567	0.9774	0.2986	0.9051	0.9048
120	0.147	0.9794	0.3014	0.9059	0.906
130	0.1406	0.9812	0.3042	0.9075	0.9071
140	0.1364	0.9827	0.3095	0.9066	0.9064
150	0.1311	0.984	0.3136	0.907	0.907
160	0.1324	0.9846	0.3048	0.906	0.9057
170	0.1308	0.9854	0.3115	0.9053	0.9051
180	0.125	0.9862	0.3126	0.9067	0.9065
190	0.1275	0.9865	0.3116	0.907	0.9069
200	0.1255	0.9867	0.3142	0.9062	0.9062
210	0.1192	0.9868	0.3138	0.9067	0.9069
220	0.1204	0.9871	0.3136	0.9074	0.9076
230	0.1215	0.9873	0.3157	0.9064	0.9062
240	0.1172	0.9879	0.316	0.9059	0.906
250	0.1213	0.9874	0.313	0.9083	0.9084
260	0.1169	0.9881	0.3186	0.9073	0.9073
270	0.1173	0.988	0.3224	0.907	0.907
280	0.1159	0.9883	0.3167	0.9077	0.9076
290	0.1187	0.9881	0.3172	0.9067	0.9063
300	0.1165	0.988	0.3178	0.907	0.9067
310	0.1171	0.9883	0.3198	0.9071	0.9071
320	0.1186	0.9881	0.315	0.9079	0.9077
330	0.1168	0.9881	0.3158	0.9083	0.9083
340	0.1185	0.9886	0.3213	0.9065	0.9063
350	0.116	0.9883	0.3163	0.907	0.9072
360	0.1181	0.9887	0.3148	0.9063	0.9066
370	0.1169	0.9885	0.3187	0.9081	0.9081
380	0.1191	0.988	0.3181	0.9062	0.9059
390	0.1161	0.9885	0.3187	0.9073	0.9071
400	0.1175	0.9885	0.3212	0.9077	0.9075
410	0.1177	0.9884	0.3186	0.9061	0.9057
420	0.118	0.9883	0.3229	0.9071	0.907
430	0.1171	0.9887	0.3199	0.9063	0.9063
440	0.1137	0.9884	0.3192	0.9073	0.9073
450	0.1163	0.9882	0.3134	0.9088	0.909
460	0.1169	0.9882	0.3209	0.9067	0.9067
470	0.1196	0.9884	0.3166	0.907	0.9068
480	0.1159	0.9888	0.3151	0.9071	0.9071
490	0.1165	0.9884	0.3154	0.9062	0.9062
500	0.1207	0.9884	0.3201	0.9068	0.9067

Plots

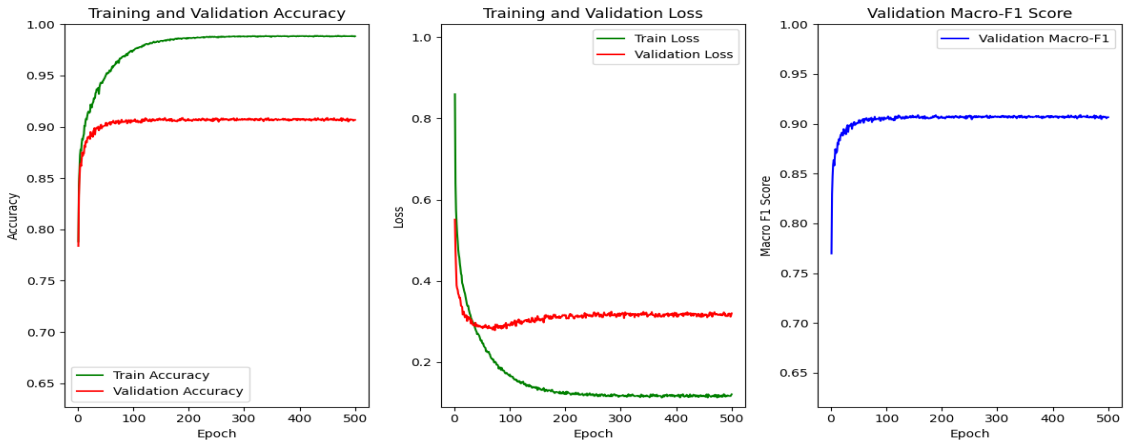

Figure 15: Various Plots

Confusion Matrix on Validation Data

Actual \ Predicted	Class 0	Class 1	Class 2	Class 3	Class 4	Class 5	Class 6	Class 7	Class 8	Class 9
Class 0	1018	0	16	35	2	1	129	0	1	0
Class 1	1	1196	2	17	1	0	1	0	1	0
Class 2	19	0	1040	5	86	0	54	0	1	0
Class 3	18	1	7	1096	34	0	23	0	4	1
Class 4	1	1	67	29	1043	0	56	0	4	1
Class 5	0	0	0	0	0	1186	0	13	0	12
Class 6	135	0	80	30	68	0	899	0	6	0
Class 7	0	0	0	0	0	20	0	1103	2	34
Class 8	3	1	3	4	7	3	11	1	1164	0
Class 9	0	0	0	0	0	10	0	27	1	1165

Table 12: Confusion Matrix

3 Best Model Analysis

3.1 Choosing the Best Model

Model 2 with `learning rate` **0.01** has the **best** `accuracy` and `macro f1 score` on the **validation dataset**.

Accuracy on Validation Data: 0.9092
Macro F1 Score on Validation Data: 0.9093

Confusion Matrix of Best Model on Validation Data

Actual \ Predicted	Class 0	Class 1	Class 2	Class 3	Class 4	Class 5	Class 6	Class 7	Class 8	Class 9
Class 0	1018	0	16	35	2	1	129	0	1	0
Class 1	1	1196	2	17	1	0	1	0	1	0
Class 2	19	0	1040	5	86	0	54	0	1	0
Class 3	18	1	7	1096	34	0	23	0	4	1
Class 4	1	1	67	29	1043	0	56	0	4	1
Class 5	0	0	0	0	0	1186	0	13	0	12
Class 6	135	0	80	30	68	0	899	0	6	0
Class 7	0	0	0	0	0	20	0	1103	2	34
Class 8	3	1	3	4	7	3	11	1	1164	0
Class 9	0	0	0	0	0	10	0	27	1	1165

Table 13: Confusion Matrix

3.2 Best Model on Test Data

Accuracy on Test Data: 0.9020
Macro F1 Score on Test Data: 0.9021

Confusion Matrix of Best Model on Test Data

Actual \ Predicted	Class 0	Class 1	Class 2	Class 3	Class 4	Class 5	Class 6	Class 7	Class 8	Class 9
Class 0	835	1	21	20	4	0	116	0	3	0
Class 1	1	981	1	12	3	0	1	0	1	0
Class 2	17	1	840	9	64	0	67	1	1	0
Class 3	19	4	11	897	40	0	25	0	3	1
Class 4	0	0	64	27	850	0	58	0	1	0
Class 5	0	0	0	0	0	974	0	17	1	8
Class 6	108	0	77	21	53	0	731	0	10	0
Class 7	0	0	0	0	0	10	0	967	0	23
Class 8	4	0	5	2	2	1	4	4	978	0
Class 9	0	0	0	0	0	6	1	26	0	967

Table 14: Confusion Matrix