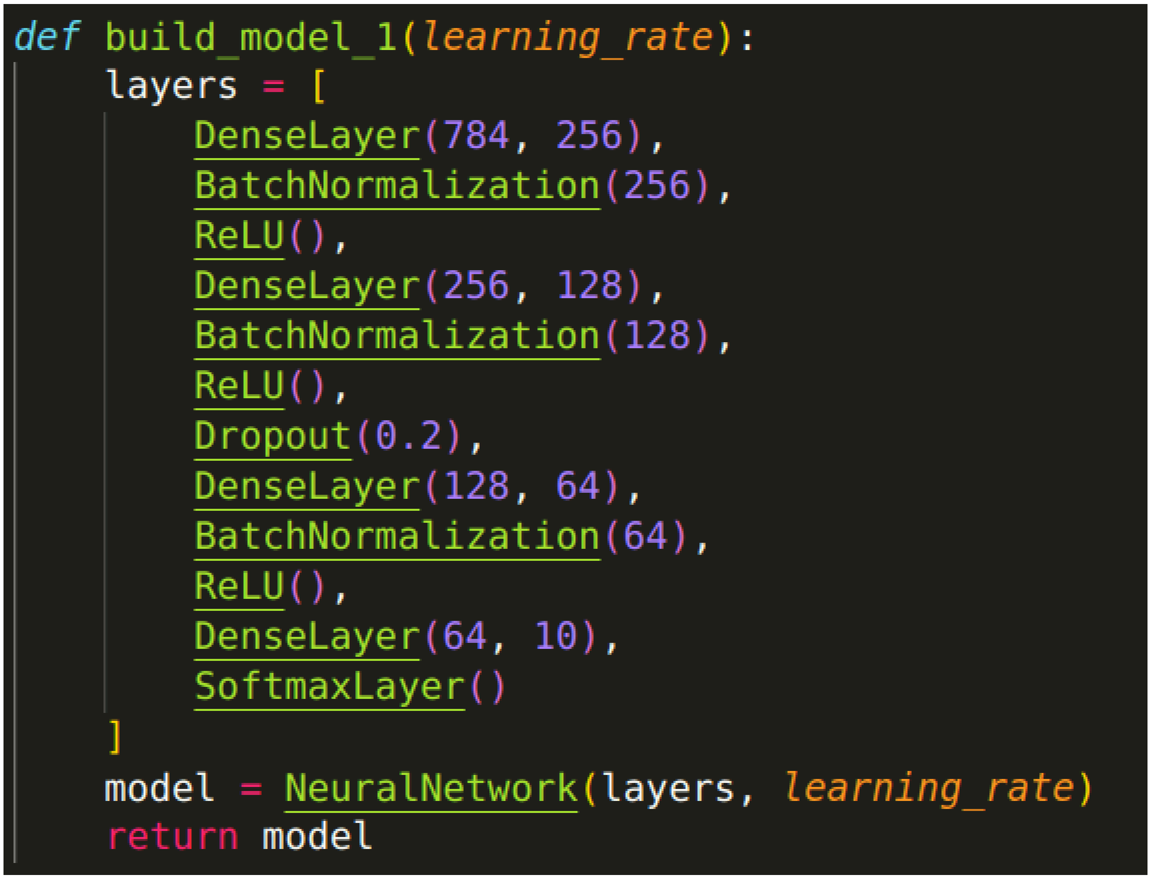
Report ML Offline-3

Name : Protoy Barai

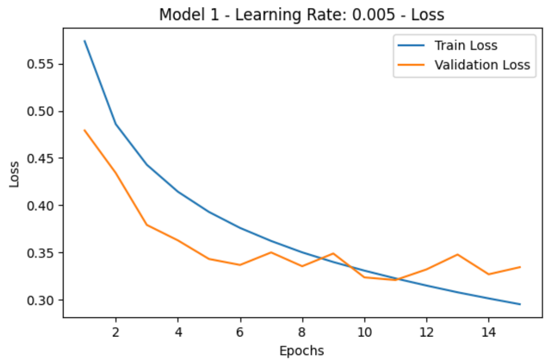
Student ID : 1905068

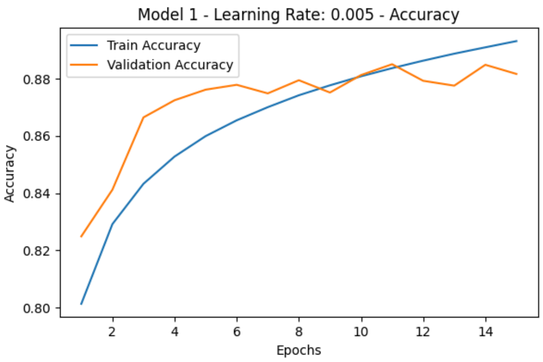
Model - 1

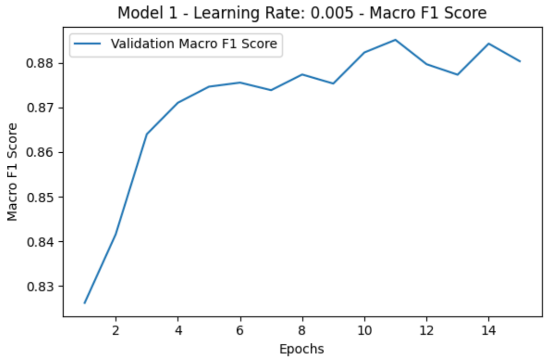


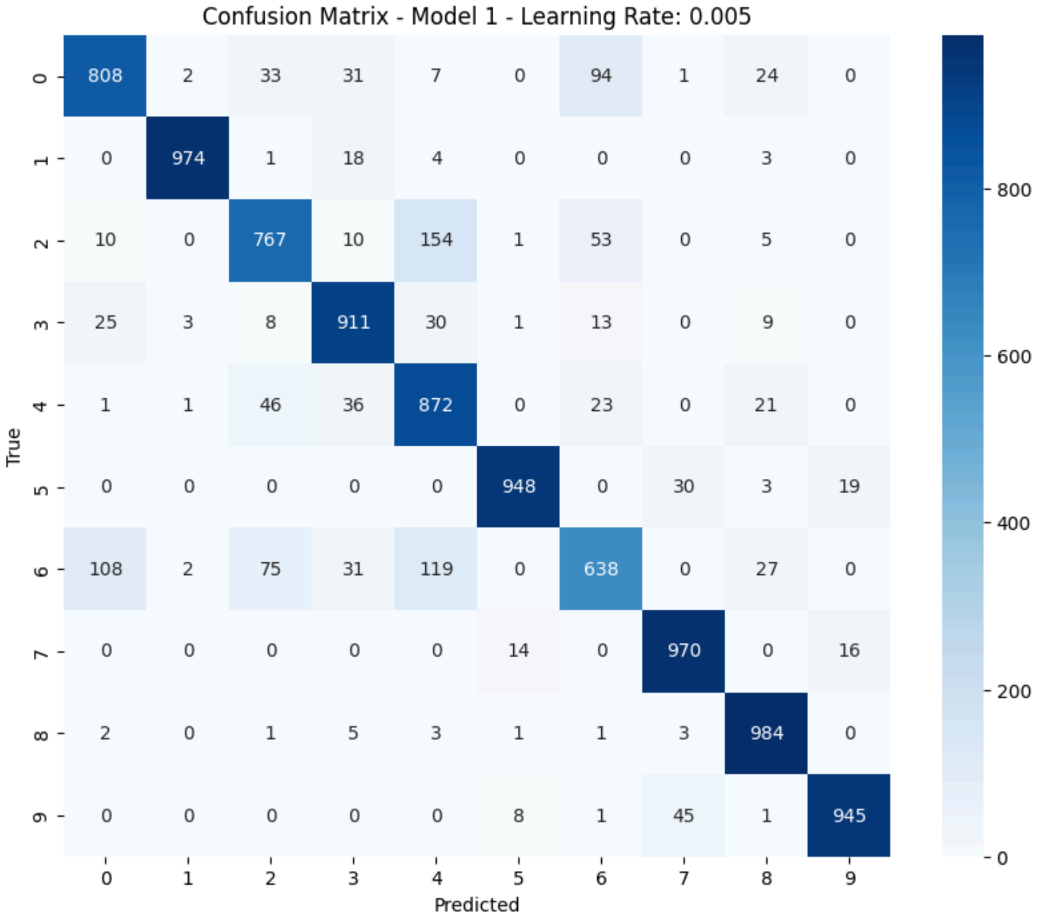
```python
def build_model_1(learning_rate):
    layers = [
        DenseLayer(784, 256),
        BatchNormalization(256),
        ReLU(),
        DenseLayer(256, 128),
        BatchNormalization(128),
        ReLU(),
        Dropout(0.2),
        DenseLayer(128, 64),
        BatchNormalization(64),
        ReLU(),
        DenseLayer(64, 10),
        SoftmaxLayer()
    ]
    model = NeuralNetwork(layers, learning_rate)
    return model
```

Model-1 with learning rate-0.005


Model 1 - Learning Rate: 0.005 - Loss

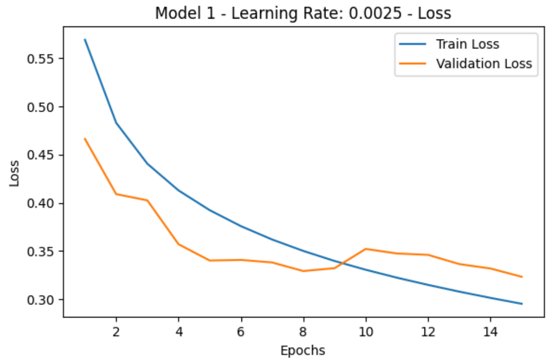

Model 1 - Learning Rate: 0.005 - Accuracy

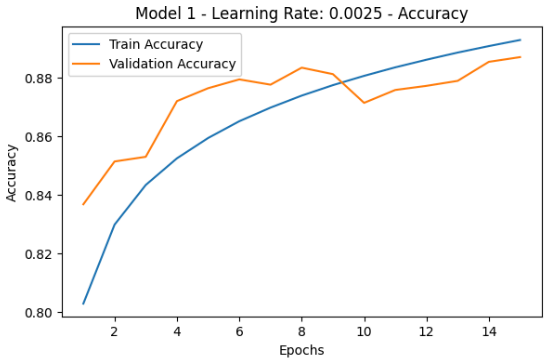

Model 1 - Learning Rate: 0.005 - Macro F1 Score

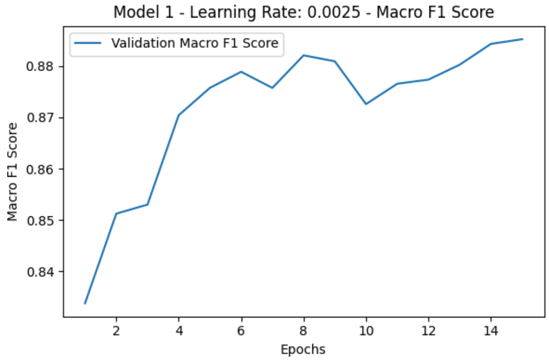

Confusion Matrix - Model 1 - Learning Rate: 0.005

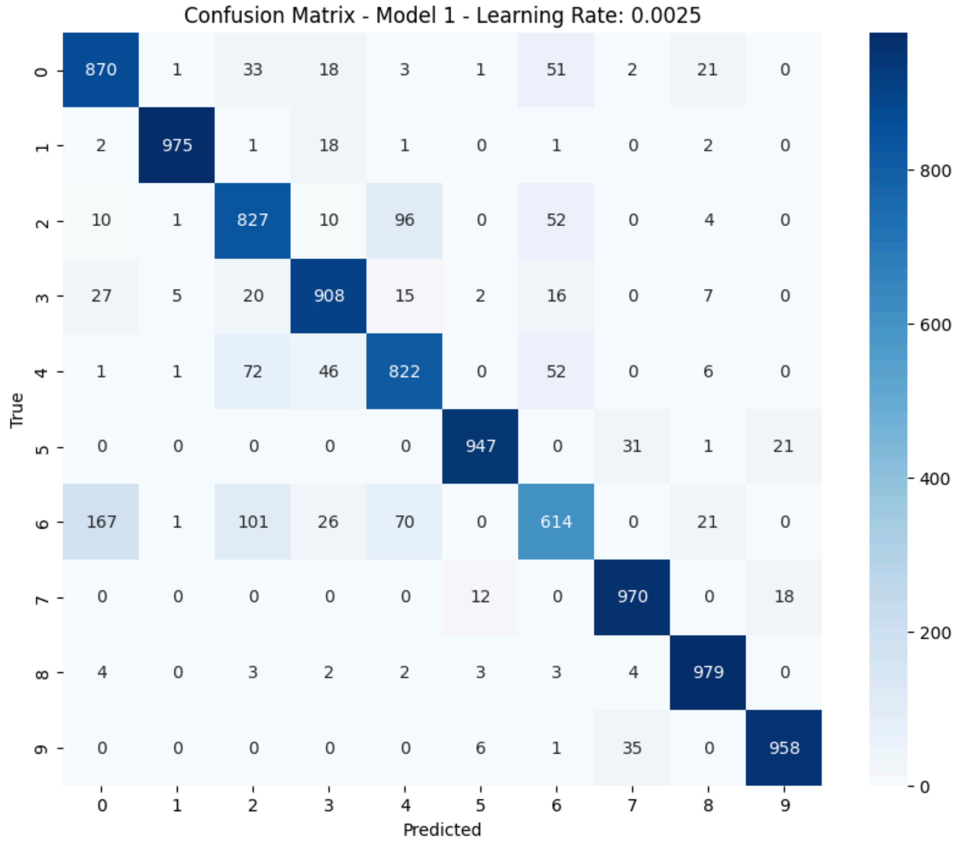
True \ Predicted	0	1	2	3	4	5	6	7	8	9
0	808	2	33	31	7	0	94	1	24	0
1	0	974	1	18	4	0	0	0	3	0
2	10	0	767	10	154	1	53	0	5	0
3	25	3	8	911	30	1	13	0	9	0
4	1	1	46	36	872	0	23	0	21	0
5	0	0	0	0	0	948	0	30	3	19
6	108	2	75	31	119	0	638	0	27	0
7	0	0	0	0	0	14	0	970	0	16
8	2	0	1	5	3	1	1	3	984	0
9	0	0	0	0	0	8	1	45	1	945

Model - 1 with learning rate 0.0025



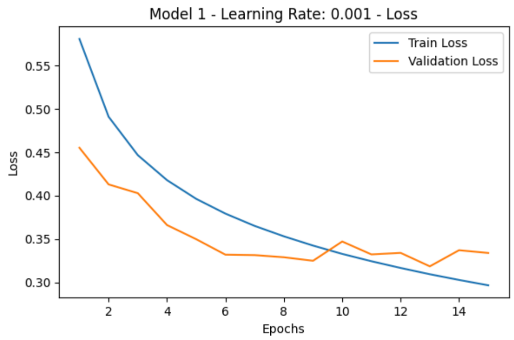


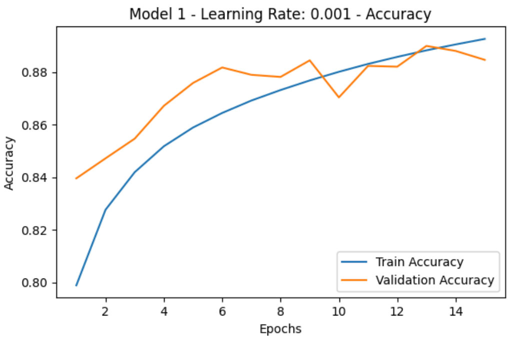


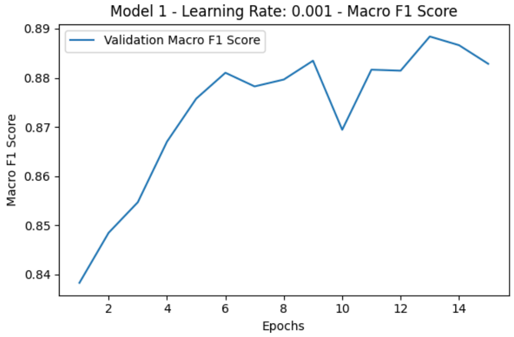


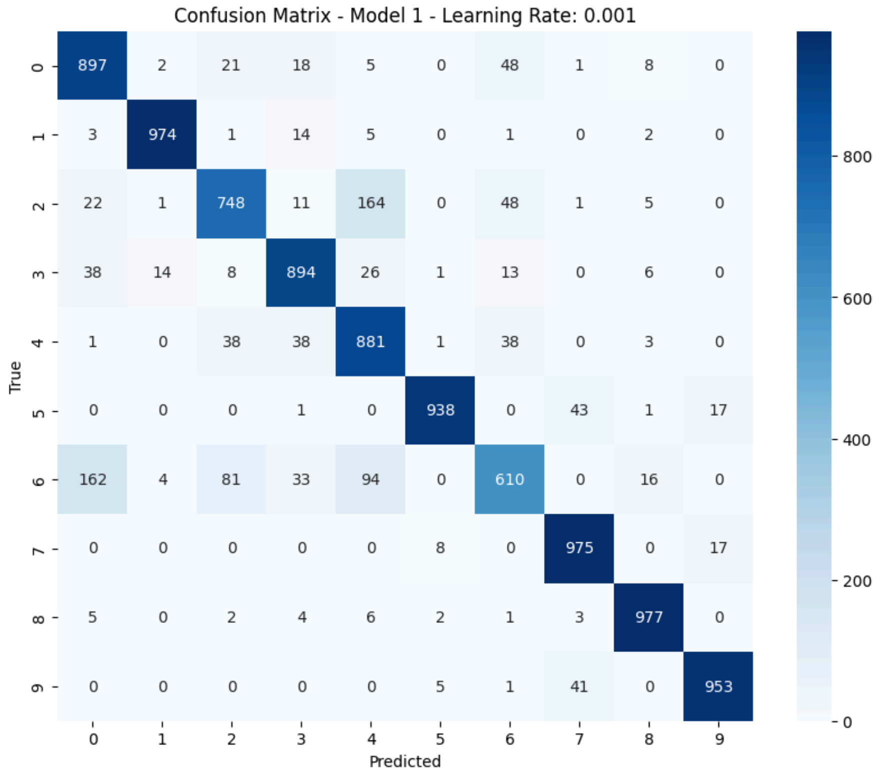
True \ Predicted	0	1	2	3	4	5	6	7	8	9
0	870	1	33	18	3	1	51	2	21	0
1	2	975	1	18	1	0	1	0	2	0
2	10	1	827	10	96	0	52	0	4	0
3	27	5	20	908	15	2	16	0	7	0
4	1	1	72	46	822	0	52	0	6	0
5	0	0	0	0	0	947	0	31	1	21
6	167	1	101	26	70	0	614	0	21	0
7	0	0	0	0	0	12	0	970	0	18
8	4	0	3	2	2	3	3	4	979	0
9	0	0	0	0	0	6	1	35	0	958

Model-1 with learning rate 0.001



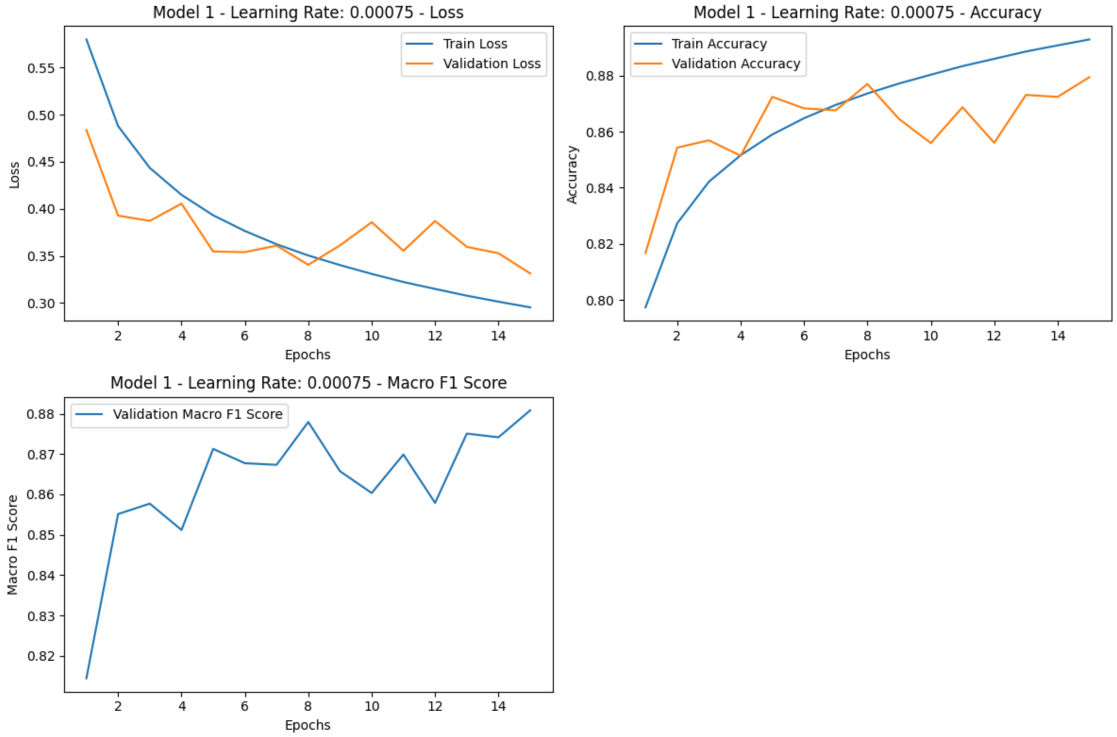






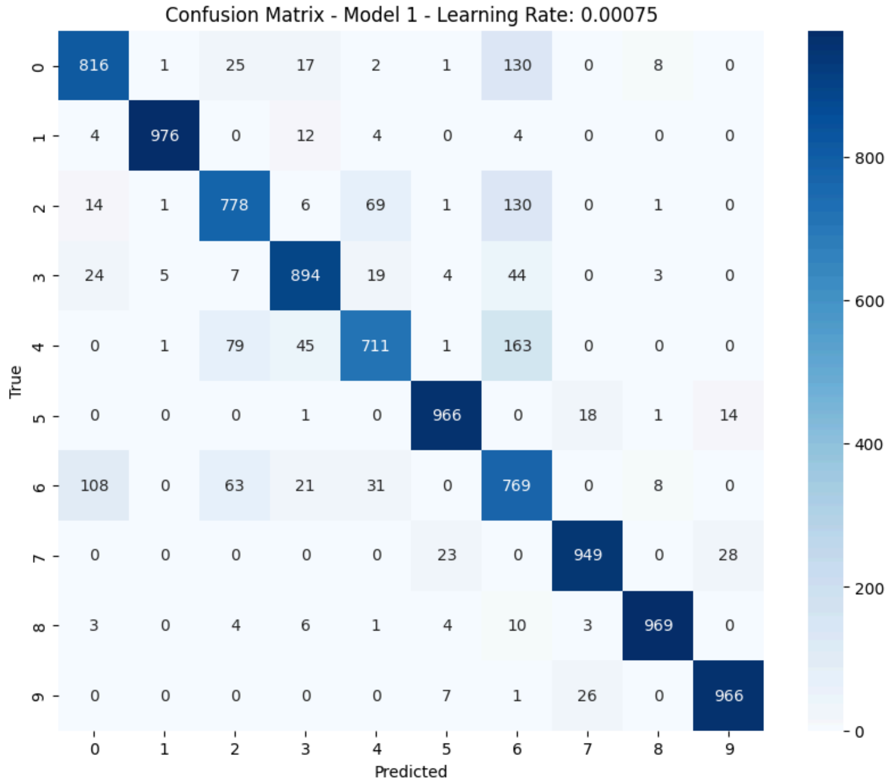
	0	1	2	3	4	5	6	7	8	9
0	897	2	21	18	5	0	48	1	8	0
1	3	974	1	14	5	0	1	0	2	0
2	22	1	748	11	164	0	48	1	5	0
3	38	14	8	894	26	1	13	0	6	0
4	1	0	38	38	881	1	38	0	3	0
5	0	0	0	1	0	938	0	43	1	17
6	162	4	81	33	94	0	610	0	16	0
7	0	0	0	0	0	8	0	975	0	17
8	5	0	2	4	6	2	1	3	977	0
9	0	0	0	0	0	5	1	41	0	953

Model-1 with learning rate 0.00075


Model 1 - Learning Rate: 0.00075 - Loss

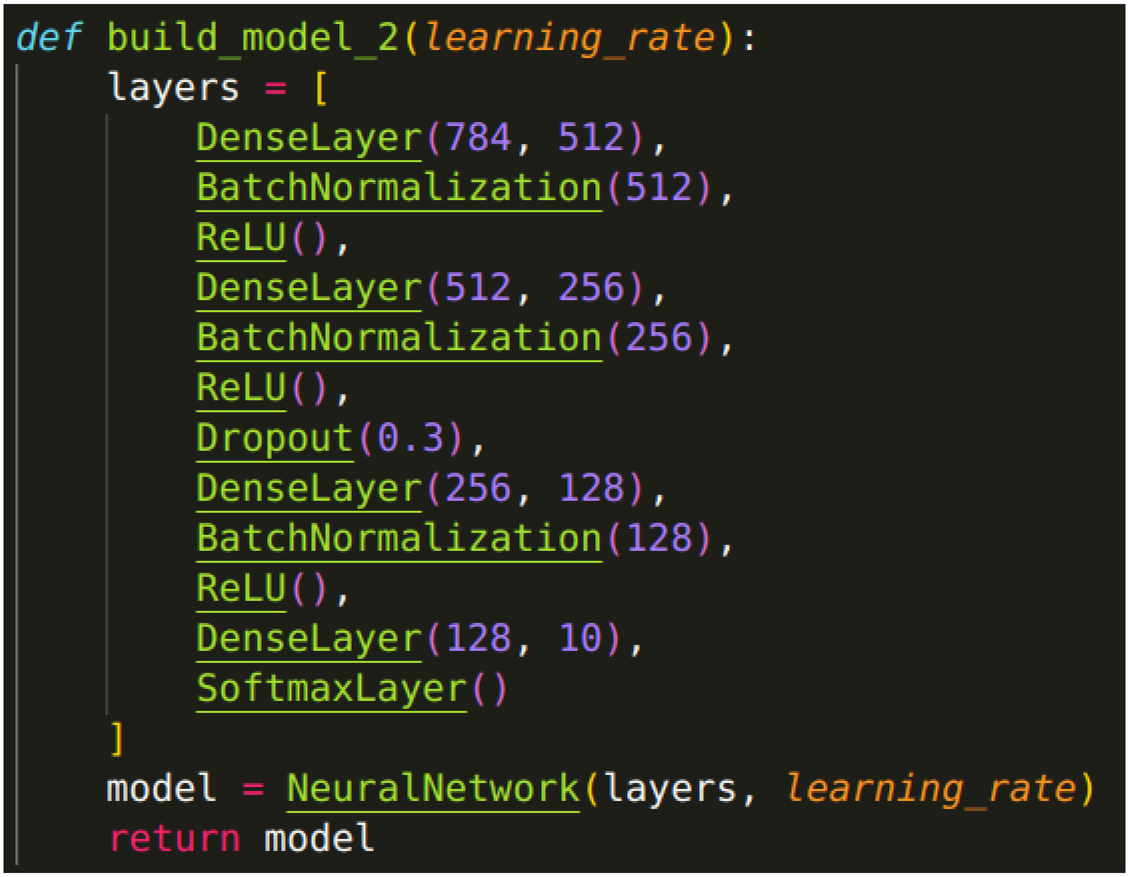
Model 1 - Learning Rate: 0.00075 - Accuracy

Model 1 - Learning Rate: 0.00075 - Macro F1 Score


Confusion Matrix - Model 1 - Learning Rate: 0.00075

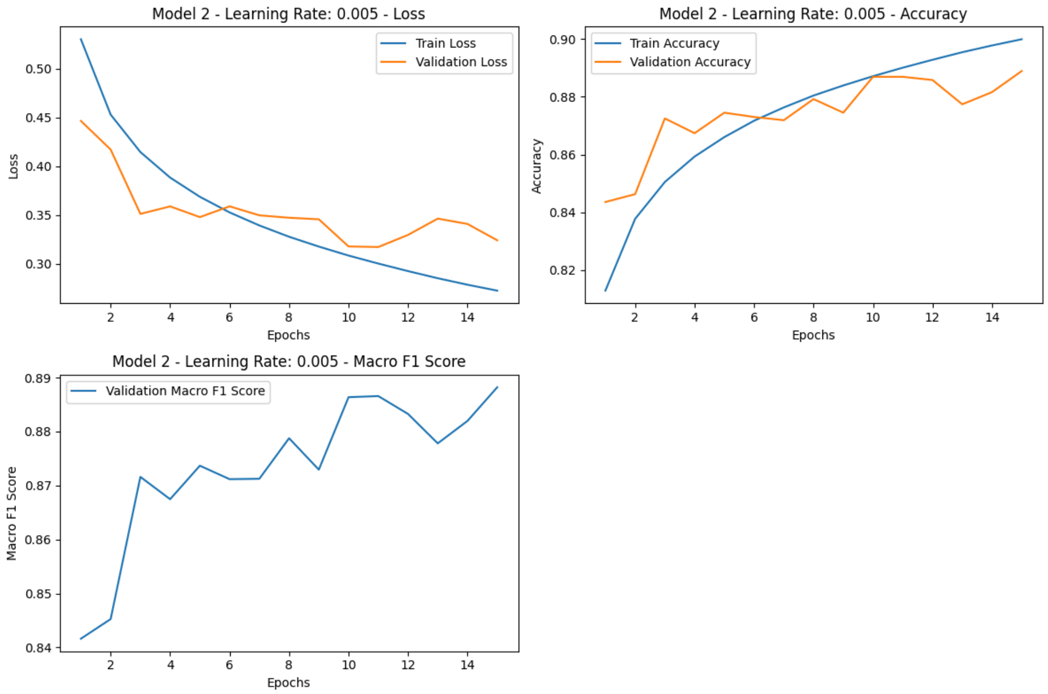
True \ Predicted	0	1	2	3	4	5	6	7	8	9
0	816	1	25	17	2	1	130	0	8	0
1	4	976	0	12	4	0	4	0	0	0
2	14	1	778	6	69	1	130	0	1	0
3	24	5	7	894	19	4	44	0	3	0
4	0	1	79	45	711	1	163	0	0	0
5	0	0	0	1	0	966	0	18	1	14
6	108	0	63	21	31	0	769	0	8	0
7	0	0	0	0	0	23	0	949	0	28
8	3	0	4	6	1	4	10	3	969	0
9	0	0	0	0	0	7	1	26	0	966

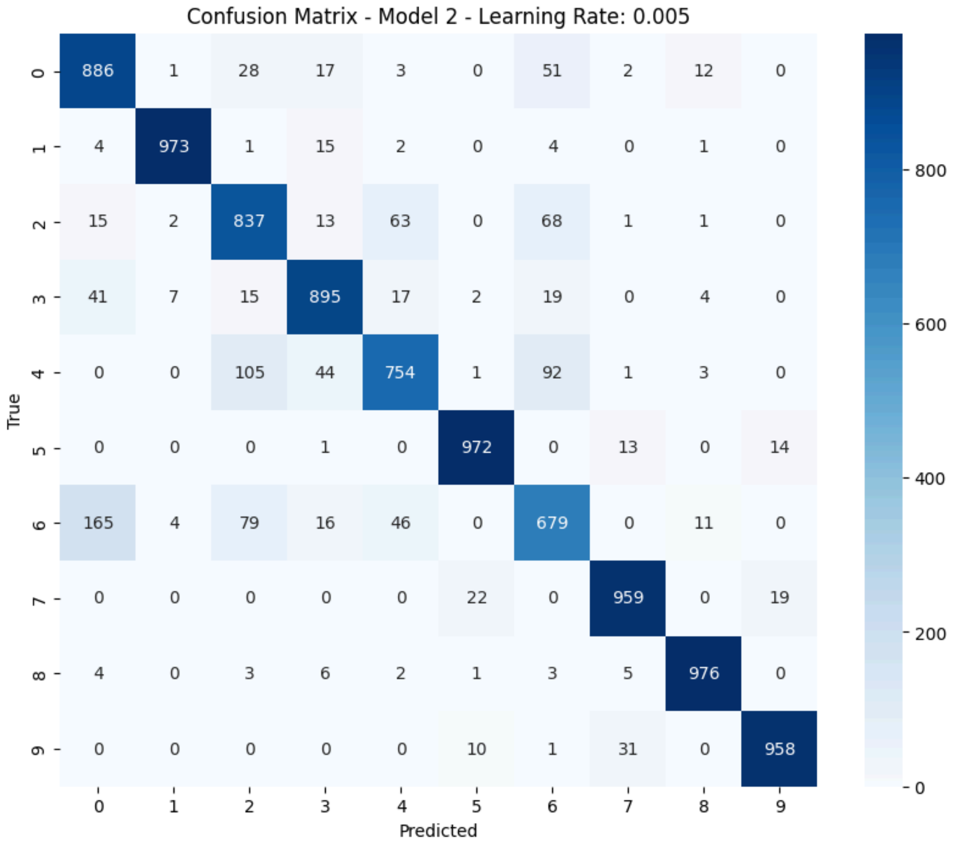
Model - 2



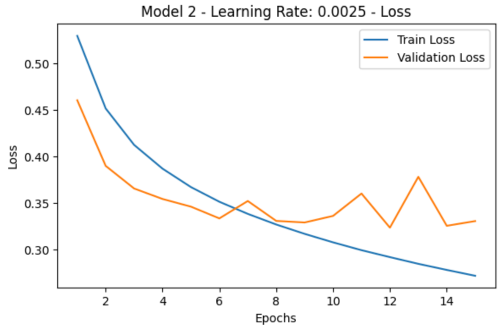
```
def build_model_2(learning_rate):
    layers = [
        DenseLayer(784, 512),
        BatchNormalization(512),
        ReLU(),
        DenseLayer(512, 256),
        BatchNormalization(256),
        ReLU(),
        Dropout(0.3),
        DenseLayer(256, 128),
        BatchNormalization(128),
        ReLU(),
        DenseLayer(128, 10),
        SoftmaxLayer()
    ]
    model = NeuralNetwork(layers, learning_rate)
    return model
```

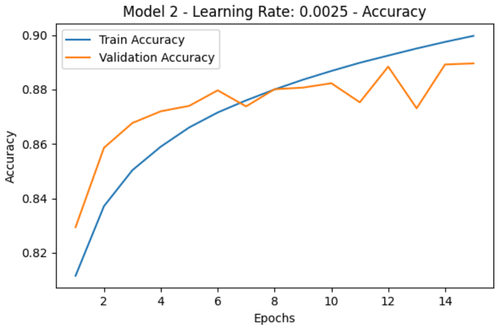
Model-2 with learning rate 0.005

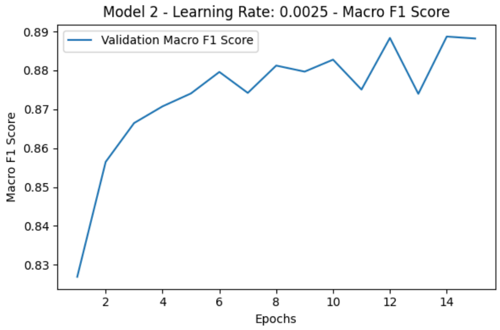

Model 2 - Learning Rate: 0.005 - Loss
Train Loss
Validation Loss
Loss
0.50
0.45
0.40
0.35
0.30
Epochs
2
4
6
8
10
12
14
Model 2 - Learning Rate: 0.005 - Accuracy
Train Accuracy
Validation Accuracy
Accuracy
0.90
0.88
0.86
0.84
0.82
Epochs
Model 2 - Learning Rate: 0.005 - Macro F1 Score
Validation Macro F1 Score
Macro F1 Score
0.89
0.88
0.87
0.86
0.85
0.84
Epochs

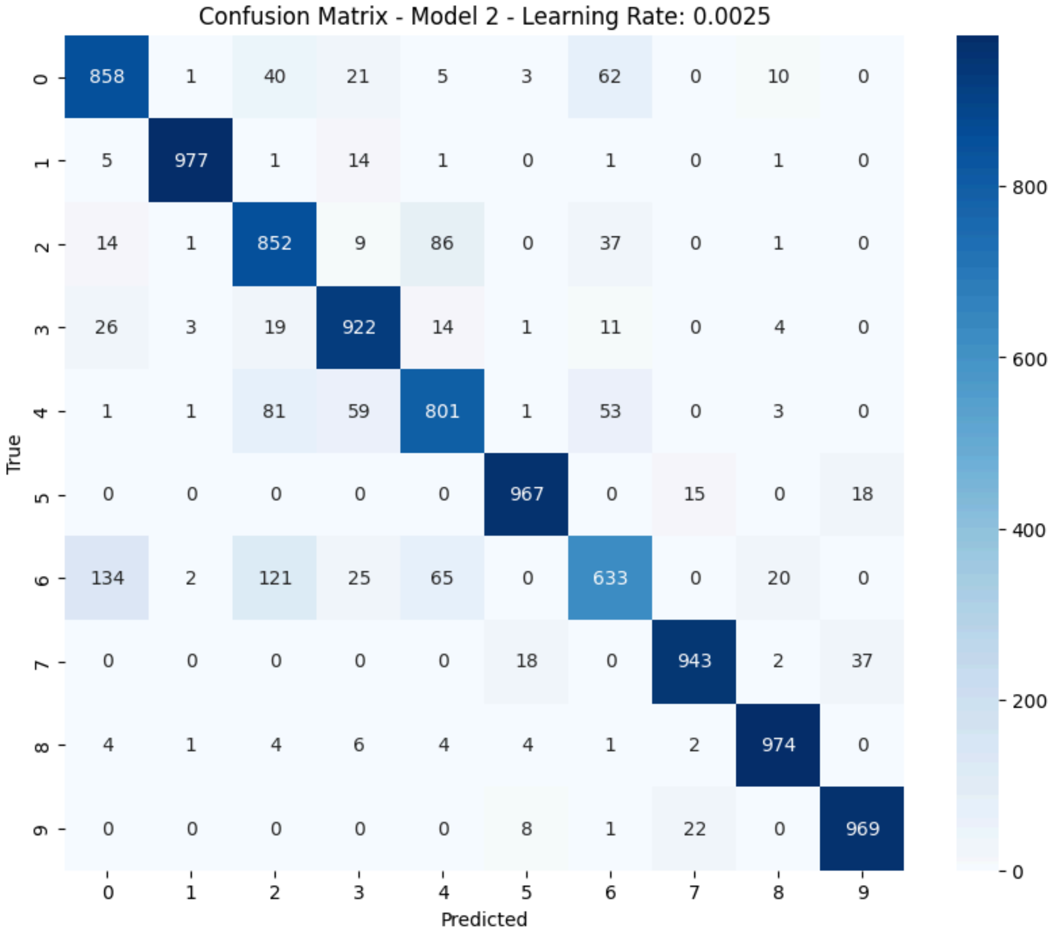

Confusion Matrix - Model 2 - Learning Rate: 0.005
True
Predicted
0 886 1 28 17 3 0 51 2 12 0
1 4 973 1 15 2 0 4 0 1 0
2 15 2 837 13 63 0 68 1 1 0
3 41 7 15 895 17 2 19 0 4 0
4 0 0 105 44 754 1 92 1 3 0
5 0 0 0 1 0 972 0 13 0 14
6 165 4 79 16 46 0 679 0 11 0
7 0 0 0 0 0 22 0 959 0 19
8 4 0 3 6 2 1 3 5 976 0
9 0 0 0 0 0 10 1 31 0 958
0 1 2 3 4 5 6 7 8 9
800
600
400
200
0

Model-2 with learning rate 0.0025



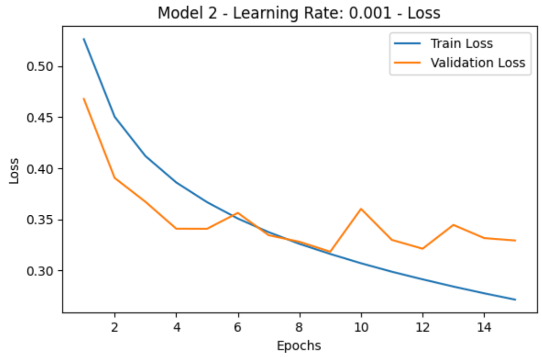




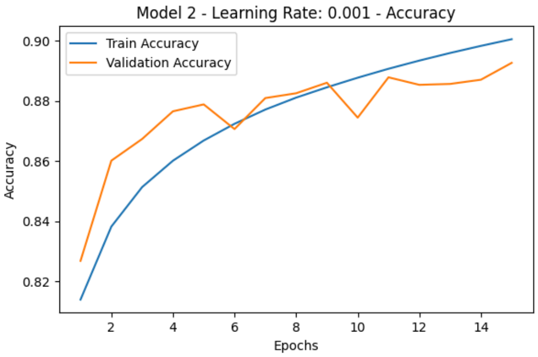


True \ Predicted	0	1	2	3	4	5	6	7	8	9
0	858	1	40	21	5	3	62	0	10	0
1	5	977	1	14	1	0	1	0	1	0
2	14	1	852	9	86	0	37	0	1	0
3	26	3	19	922	14	1	11	0	4	0
4	1	1	81	59	801	1	53	0	3	0
5	0	0	0	0	0	967	0	15	0	18
6	134	2	121	25	65	0	633	0	20	0
7	0	0	0	0	0	18	0	943	2	37
8	4	1	4	6	4	4	1	2	974	0
9	0	0	0	0	0	8	1	22	0	969

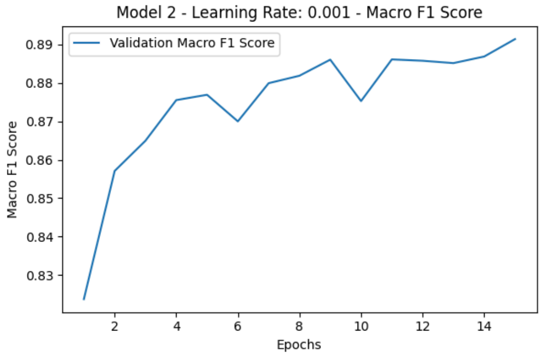
Model-2 with learning rate 0.001



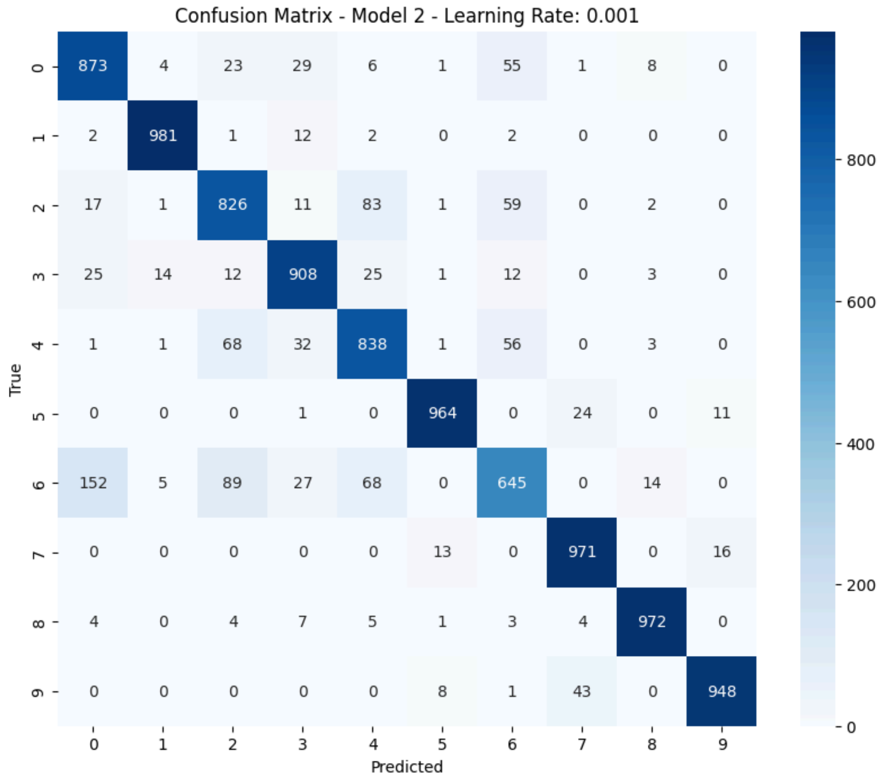
Model 2 - Learning Rate: 0.001 - Loss



Model 2 - Learning Rate: 0.001 - Accuracy



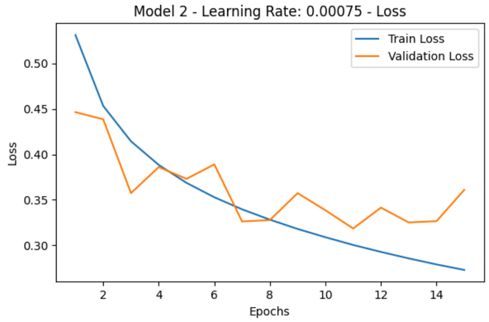
Model 2 - Learning Rate: 0.001 - Macro F1 Score

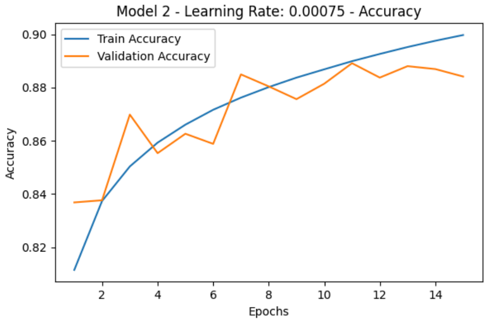


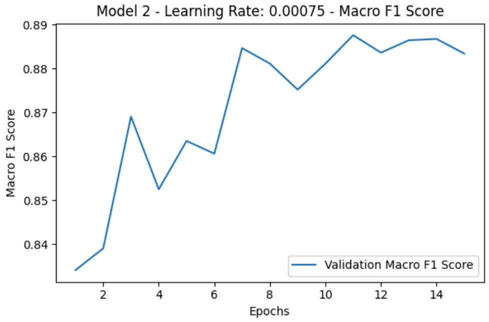
Confusion Matrix - Model 2 - Learning Rate: 0.001

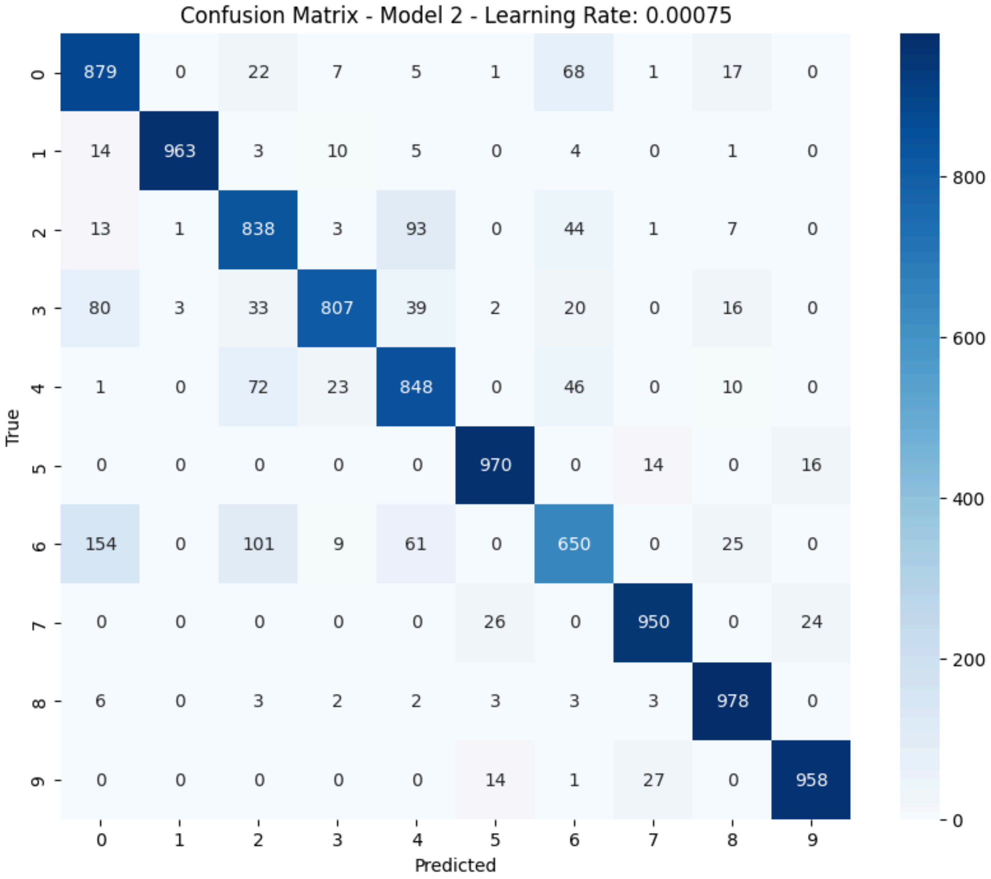
True \ Predicted	0	1	2	3	4	5	6	7	8	9
0	873	4	23	29	6	1	55	1	8	0
1	2	981	1	12	2	0	2	0	0	0
2	17	1	826	11	83	1	59	0	2	0
3	25	14	12	908	25	1	12	0	3	0
4	1	1	68	32	838	1	56	0	3	0
5	0	0	0	1	0	964	0	24	0	11
6	152	5	89	27	68	0	645	0	14	0
7	0	0	0	0	0	13	0	971	0	16
8	4	0	4	7	5	1	3	4	972	0
9	0	0	0	0	0	8	1	43	0	948

Model-2 with learning rate 0.00075



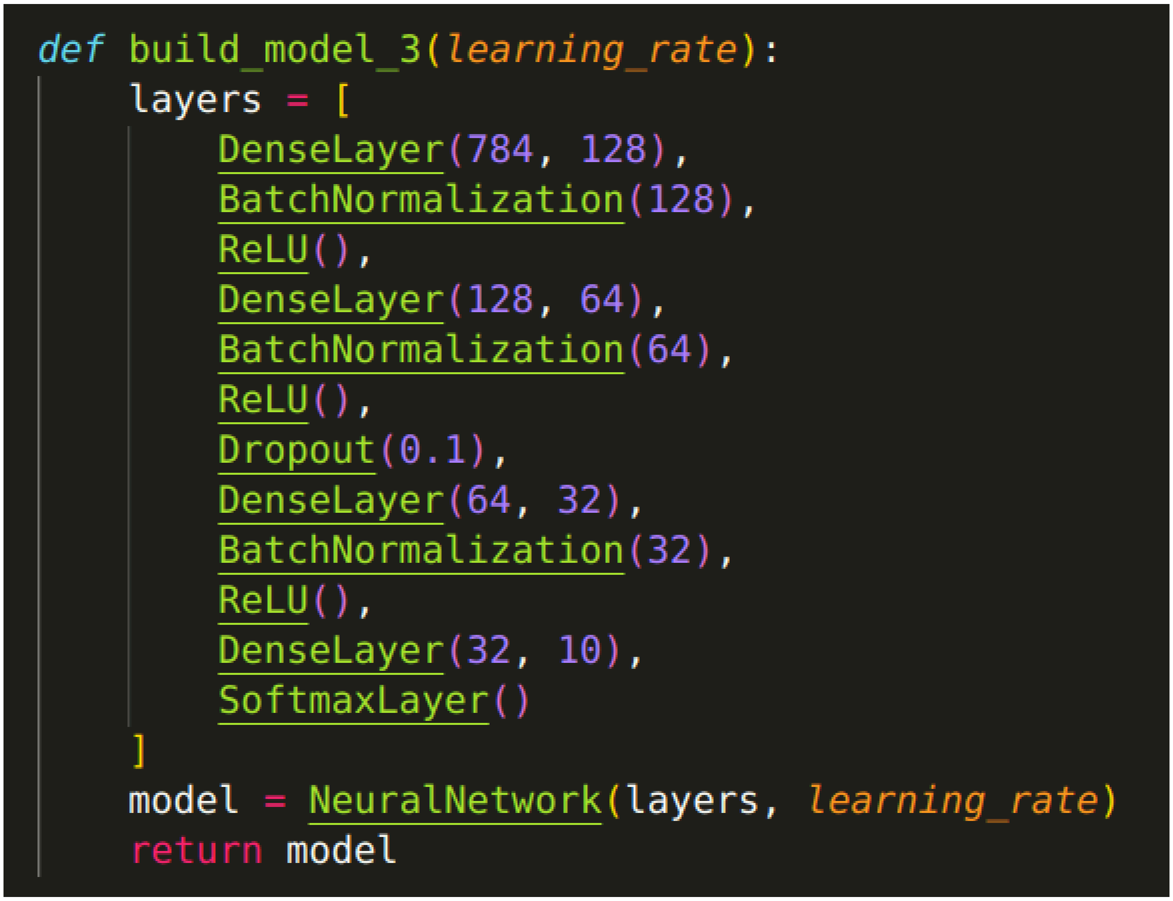





Confusion Matrix - Model 2 - Learning Rate: 0.00075

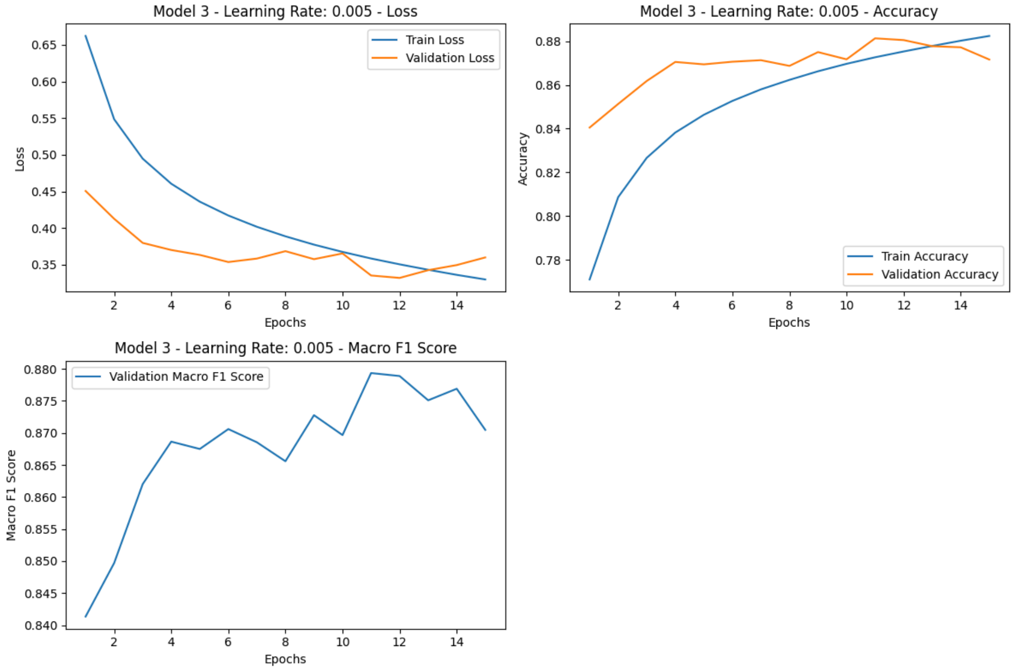
True \ Predicted	0	1	2	3	4	5	6	7	8	9
0	879	0	22	7	5	1	68	1	17	0
1	14	963	3	10	5	0	4	0	1	0
2	13	1	838	3	93	0	44	1	7	0
3	80	3	33	807	39	2	20	0	16	0
4	1	0	72	23	848	0	46	0	10	0
5	0	0	0	0	0	970	0	14	0	16
6	154	0	101	9	61	0	650	0	25	0
7	0	0	0	0	0	26	0	950	0	24
8	6	0	3	2	2	3	3	3	978	0
9	0	0	0	0	0	14	1	27	0	958

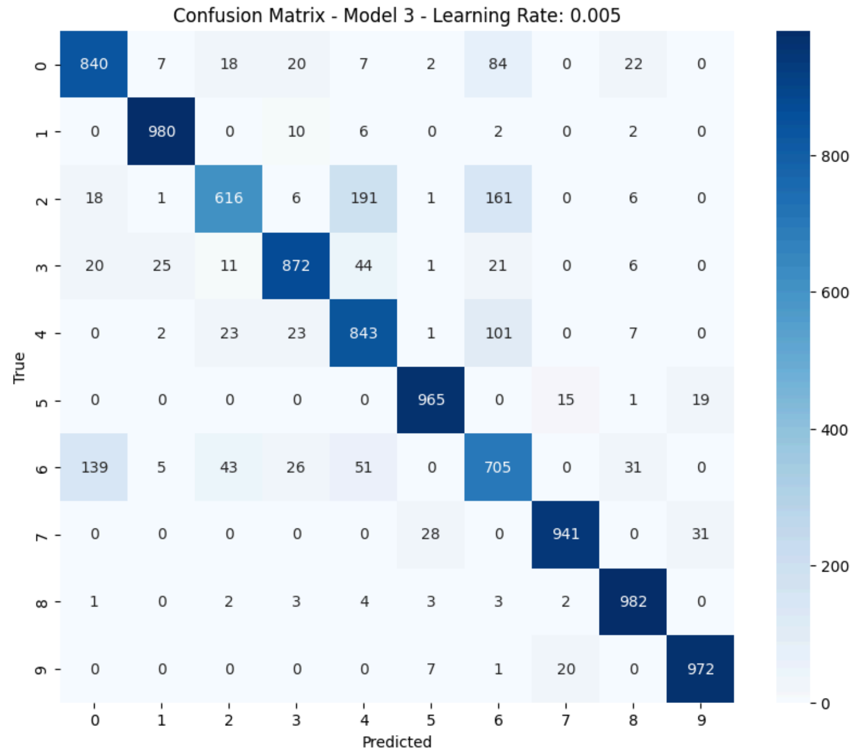
Model-3



```
def build_model_3(learning_rate):
    layers = [
        DenseLayer(784, 128),
        BatchNormalization(128),
        ReLU(),
        DenseLayer(128, 64),
        BatchNormalization(64),
        ReLU(),
        Dropout(0.1),
        DenseLayer(64, 32),
        BatchNormalization(32),
        ReLU(),
        DenseLayer(32, 10),
        SoftmaxLayer()
    ]
    model = NeuralNetwork(layers, learning_rate)
    return model
```

Model-3 with learning rate 0.005

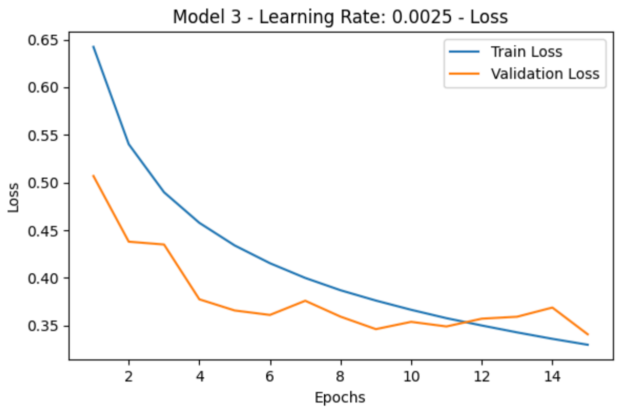


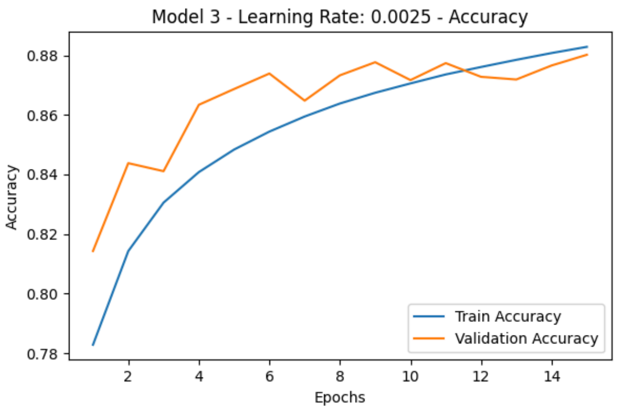


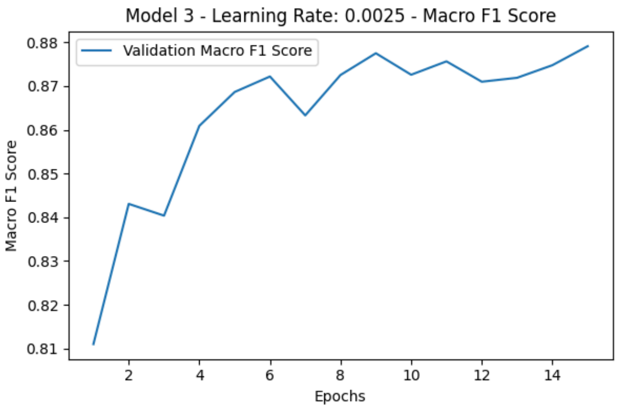
Confusion Matrix - Model 3 - Learning Rate: 0.005

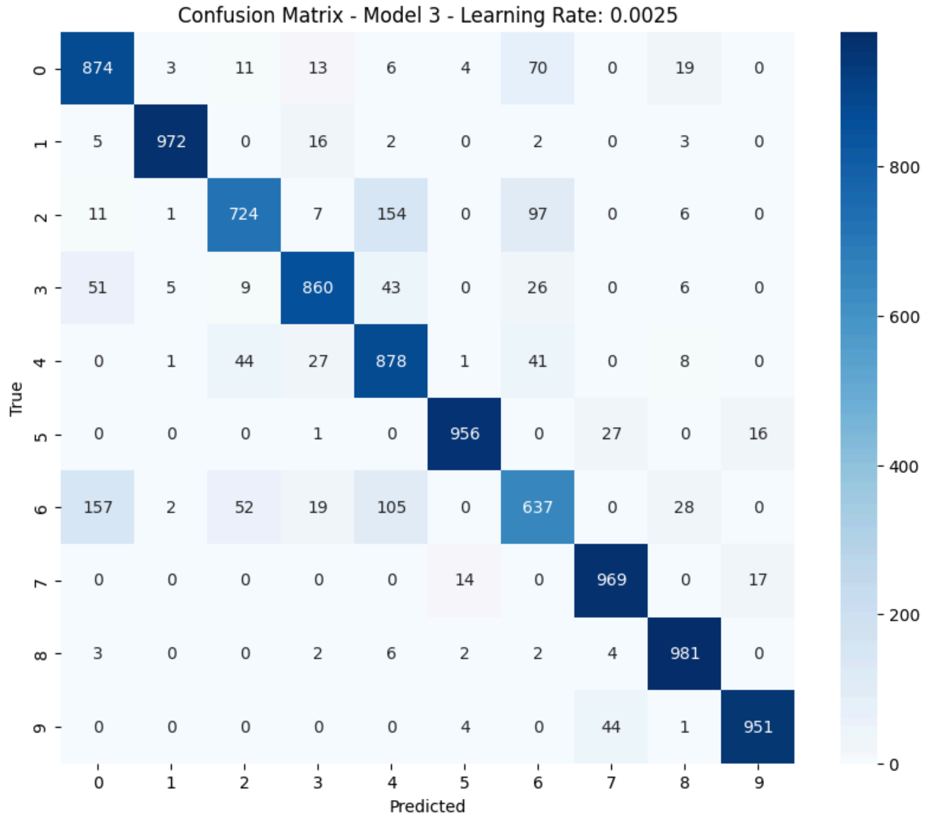
	0	1	2	3	4	5	6	7	8	9
0	840	7	18	20	7	2	84	0	22	0
1	0	980	0	10	6	0	2	0	2	0
2	18	1	616	6	191	1	161	0	6	0
3	20	25	11	872	44	1	21	0	6	0
4	0	2	23	23	843	1	101	0	7	0
5	0	0	0	0	0	965	0	15	1	19
6	139	5	43	26	51	0	705	0	31	0
7	0	0	0	0	0	28	0	941	0	31
8	1	0	2	3	4	3	3	2	982	0
9	0	0	0	0	0	7	1	20	0	972

Model-3 with learning rate 0.0025

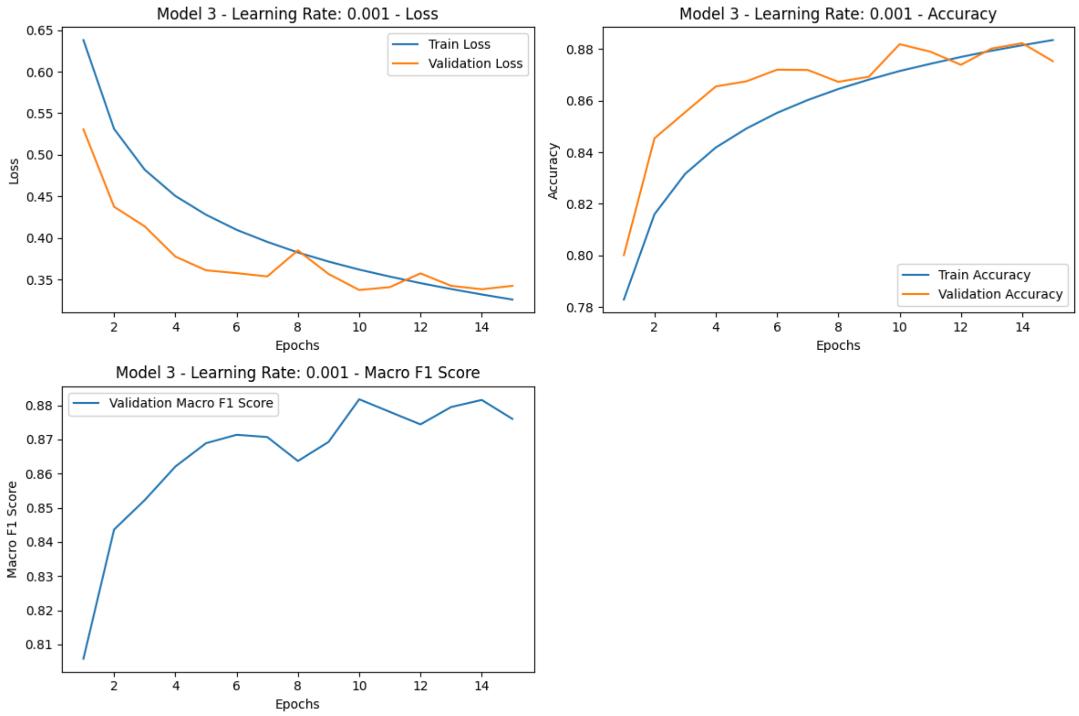

Model 3 - Learning Rate: 0.0025 - Loss
Train Loss
Validation Loss
Loss
0.65
0.60
0.55
0.50
0.45
0.40
0.35
2
4
6
8
10
12
14
Epochs

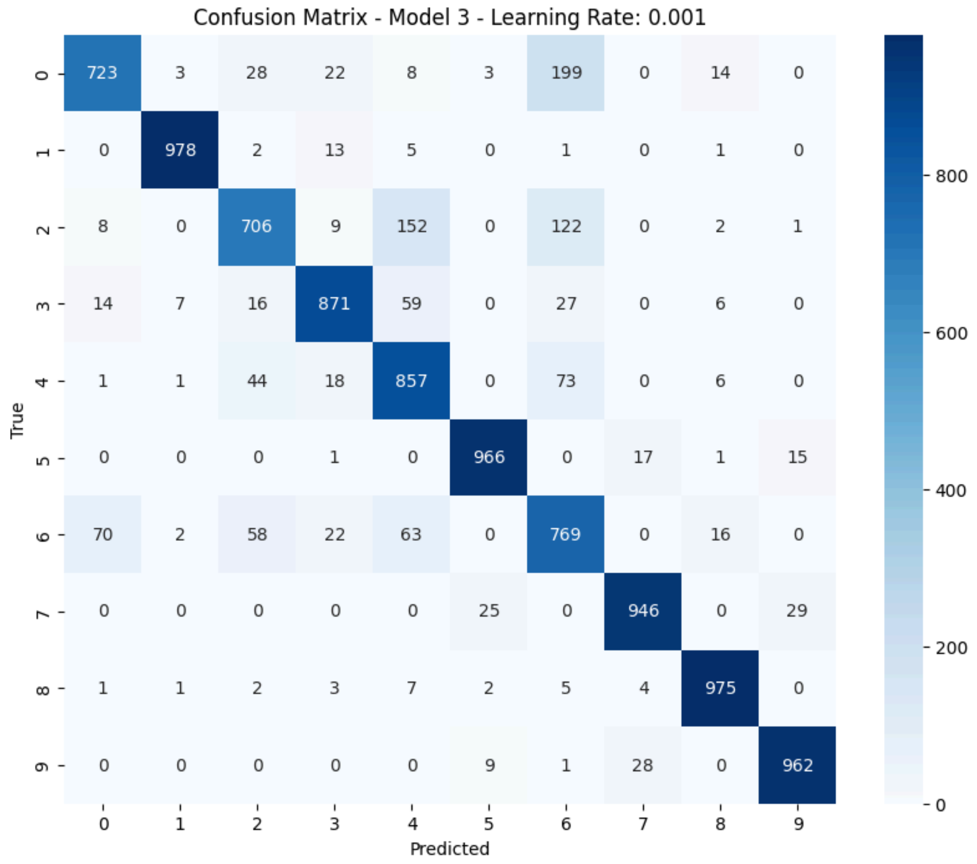

Model 3 - Learning Rate: 0.0025 - Accuracy
Accuracy
0.88
0.86
0.84
0.82
0.80
0.78
Train Accuracy
Validation Accuracy
2
4
6
8
10
12
14
Epochs


Model 3 - Learning Rate: 0.0025 - Macro F1 Score
Validation Macro F1 Score
Macro F1 Score
0.88
0.87
0.86
0.85
0.84
0.83
0.82
0.81
2
4
6
8
10
12
14
Epochs


Confusion Matrix - Model 3 - Learning Rate: 0.0025
True
Predicted
0 874 3 11 13 6 4 70 0 19 0
1 5 972 0 16 2 0 2 0 3 0
2 11 1 724 7 154 0 97 0 6 0
3 51 5 9 860 43 0 26 0 6 0
4 0 1 44 27 878 1 41 0 8 0
5 0 0 0 1 0 956 0 27 0 16
6 157 2 52 19 105 0 637 0 28 0
7 0 0 0 0 0 14 0 969 0 17
8 3 0 0 2 6 2 2 4 981 0
9 0 0 0 0 0 4 0 44 1 951
0 1 2 3 4 5 6 7 8 9
800
600
400
200
0

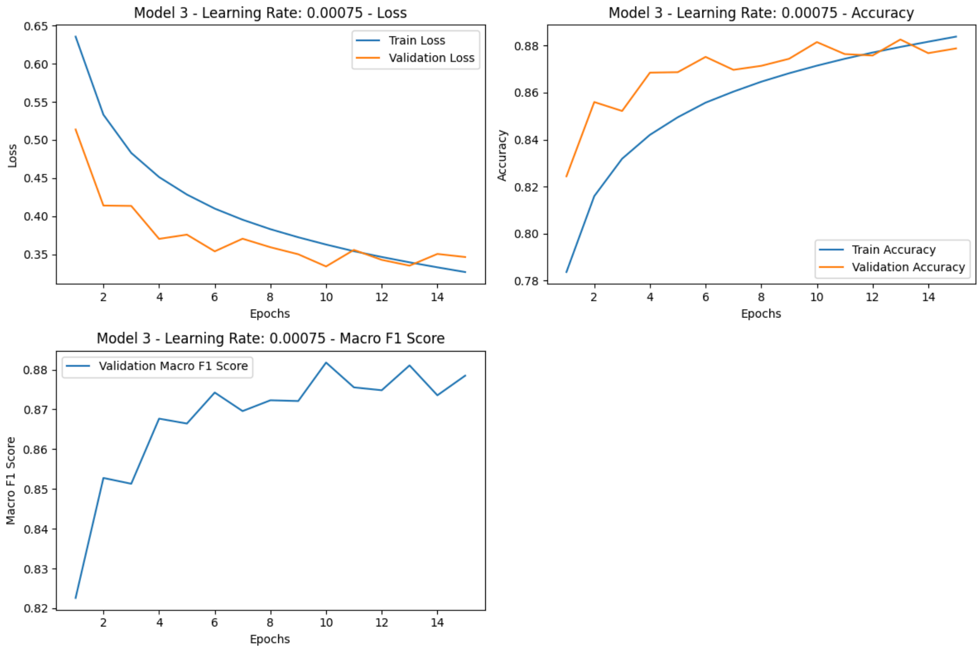
Model-3 with learning rate 0.001





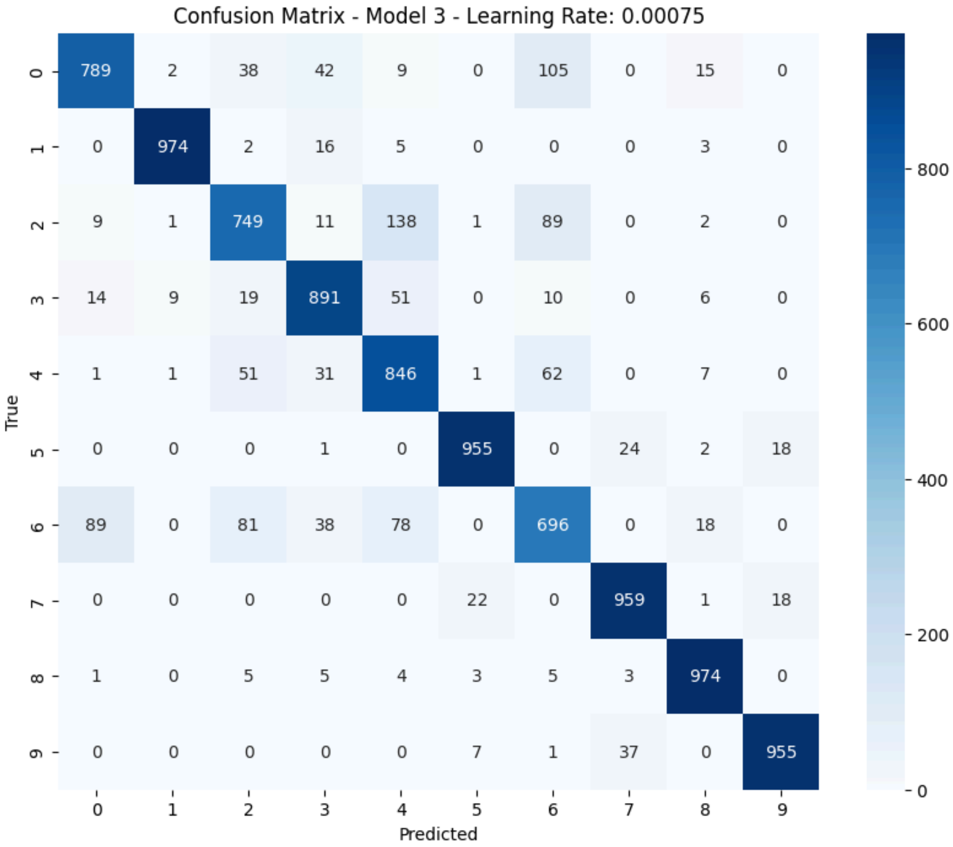
True \ Predicted	0	1	2	3	4	5	6	7	8	9
0	723	3	28	22	8	3	199	0	14	0
1	0	978	2	13	5	0	1	0	1	0
2	8	0	706	9	152	0	122	0	2	1
3	14	7	16	871	59	0	27	0	6	0
4	1	1	44	18	857	0	73	0	6	0
5	0	0	0	1	0	966	0	17	1	15
6	70	2	58	22	63	0	769	0	16	0
7	0	0	0	0	0	25	0	946	0	29
8	1	1	2	3	7	2	5	4	975	0
9	0	0	0	0	0	9	1	28	0	962

Model-3 with learning rate 0.00075


Model 3 - Learning Rate: 0.00075 - Loss

Model 3 - Learning Rate: 0.00075 - Accuracy

Model 3 - Learning Rate: 0.00075 - Macro F1 Score


Confusion Matrix - Model 3 - Learning Rate: 0.00075

True \ Predicted	0	1	2	3	4	5	6	7	8	9
0	789	2	38	42	9	0	105	0	15	0
1	0	974	2	16	5	0	0	0	3	0
2	9	1	749	11	138	1	89	0	2	0
3	14	9	19	891	51	0	10	0	6	0
4	1	1	51	31	846	1	62	0	7	0
5	0	0	0	1	0	955	0	24	2	18
6	89	0	81	38	78	0	696	0	18	0
7	0	0	0	0	0	22	0	959	1	18
8	1	0	5	5	4	3	5	3	974	0
9	0	0	0	0	0	7	1	37	0	955

Model Performance :

Model	Learning Rate	Training Loss	Validation Loss	Train Accuracy	Validation Accuracy	Validation macro F1
Model 2	**0.00100**	**0.271397**	**0.329285**	**0.900464**	**0.8926**	**0.891347**
Model 2	0.00250	0.271912	0.330645	0.899719	0.8896	0.888197
Model 2	0.00500	0.272389	0.324055	0.899901	0.8889	0.888195
Model 1	0.00250	0.295305	0.323335	0.892844	0.8870	0.885265
Model 2	0.00075	0.272682	0.360921	0.899693	0.8841	0.883375
Model 1	0.00100	0.296605	0.333930	0.892667	0.8847	0.882844
Model 1	0.00075	0.295227	0.331307	0.892887	0.8794	0.880822
Model 1	0.00500	0.295134	0.334354	0.893186	0.8817	0.880303
Model 3	0.00250	0.329868	0.340848	0.882859	0.8802	0.879105
Model 3	0.00075	0.326605	0.346331	0.883825	0.8788	0.878500
Model 3	0.00100	0.325852	0.342397	0.883494	0.8788	0.878500
Model 3	0.00500	0.329797	0.359857	0.882390	0.8716	0.870483

Best Model :
Model 2 with learning rate 0.001f