COURSE NAME: Hillel - Machine Learning. Lesson 10 (30.03.2023 19:15,) 10. Metrics, vanishing / exploding gradient, activation functions
LECTURER: Yevhen Krasnokutsky
STUDENT: Maryna Lysenko

HW 9 REPORT "Classification metrics on MNIST".

- **Accuracy (per class and general)**
- **Precision (per class and general)**
- **Recall (per class and general)**
- **F1-score (per class and general)**
- **Confusion matrix**
- **Classification report**

Model description

MnistMlp(

(wih): Linear(in_features=784, out_features=200, bias=True)

(bn1): BatchNorm1d(200, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)

(dropout1): Dropout(p=0.5, inplace=False)

(hidden2): Linear(in_features=200, out_features=150, bias=True)

(bn2): BatchNorm1d(150, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)

(dropout2): Dropout(p=0.3, inplace=False)

(hidden3): Linear(in_features=150, out_features=50, bias=True)

(bn3): BatchNorm1d(50, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)

(dropout3): Dropout(p=0.2, inplace=False)

(who): Linear(in_features=50, out_features=10, bias=True)

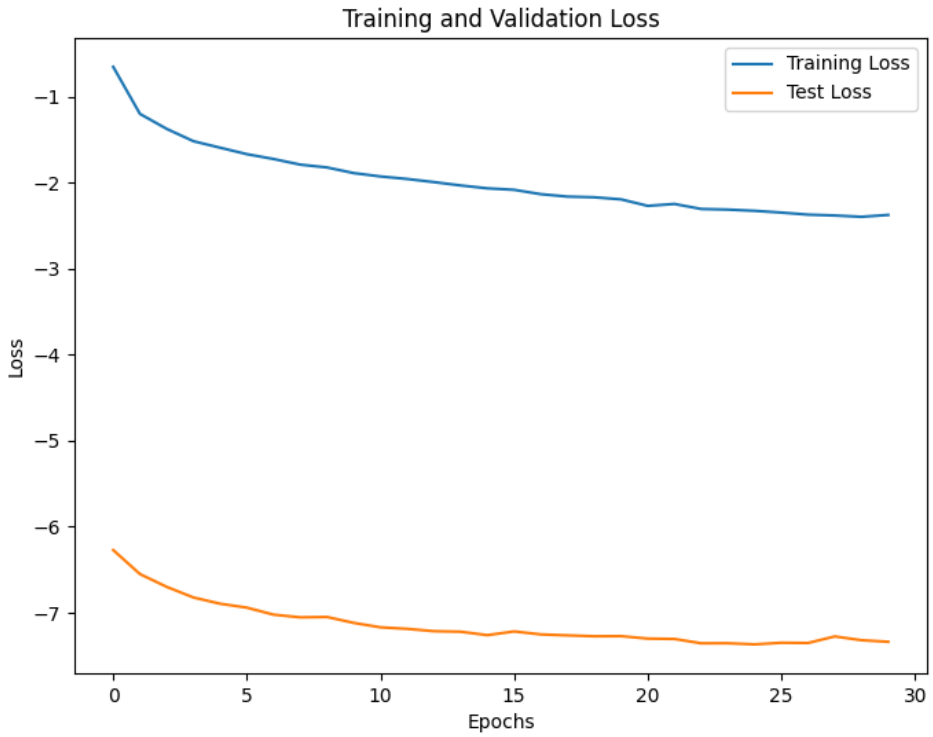
(bn4): BatchNorm1d(10, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)

(activation): **ReLU**()

)

- learning_rate = 0.1
- batch_size = 100
- epochs = 30

Test set: **Average loss**: 0.0007, **Accuracy**: 9821/10000 (98%)



Test data class sizes

Class	0	1	2	3	4	5	6	7	8	9
size	980	1135	1032	1010	982	892	958	1028	974	1009

The dataset is pretty balanced with classes ranging from 892 to 1135 samples.

Classification metrics

General:

accuracy	precision	recall	F1-score
0.9821	0.9821	0.982	0.982

The model achieved high accuracy and other performance metrics, with an overall accuracy of 0.9821 and an F1-score of 0.9820.

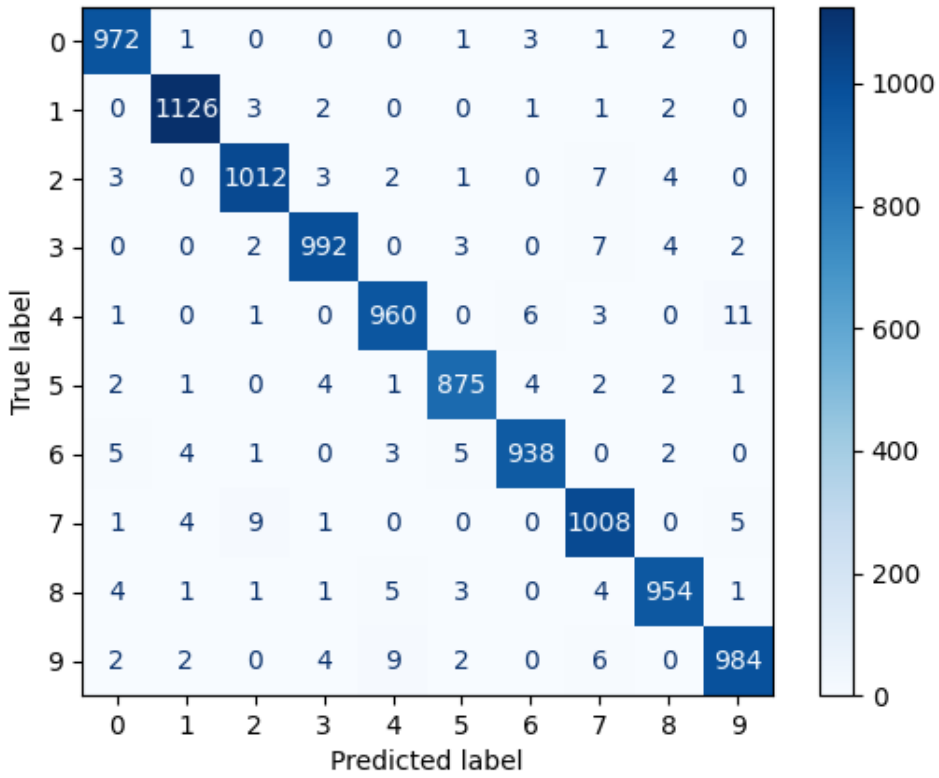
Per class :

Class	Accuracy	Precision	Recall	F1-score
0	0.9918	0.9818	0.9918	0.9868
1	0.9921	0.9886	0.9921	0.9903
2	0.9806	0.9835	0.9806	0.982
3	0.9822	0.9851	0.9822	0.9836
4	0.9776	0.9796	0.9776	0.9786
5	0.9809	0.9831	0.9809	0.982
6	0.9791	0.9853	0.9791	0.9822
7	0.9805	0.9702	0.9805	0.9753
8	0.9795	0.9835	0.9795	0.9815
9	0.9752	0.9801	0.9752	0.9776

The model has high precision and recall for most of the classes, indicating that it is able to both identify and correctly classify the majority of the samples.

Class 0, 1 have the highest accuracy and F1 scores, while class 7, 9 has the lowest.

Confusion matrix:



The confusion matrix shows that the model has problems distinguishing between classes 4 and 9

	precision	recall	f1-score	support
0	0.98	0.99	0.99	980
1	0.99	0.99	0.99	1135
2	0.98	0.98	0.98	1032
3	0.99	0.98	0.98	1010
4	0.98	0.98	0.98	982
5	0.98	0.98	0.98	892
6	0.99	0.98	0.98	958
7	0.97	0.98	0.98	1028
8	0.98	0.98	0.98	974
9	0.98	0.98	0.98	1009
accuracy			0.98	10000
macro avg	0.98	0.98	0.98	10000
weighted avg	0.98	0.98	0.98	10000

Conclusions:

1. The model achieved high accuracy and other performance metrics, with an overall accuracy of 0.9821 and an F1-score of 0.9820.
2. Class 0, 1 have the highest accuracy and F1 scores, while class 7, 9 has the lowest.

3. The model has high precision and recall for most of the classes, indicating that it is able to both identify and correctly classify the majority of the samples.
4. The confusion matrix shows that the model has problems with distinguishing between classes 4 and 9