

# **VMMT3 Projekt**

## **MNIST dataset**

# Obsah

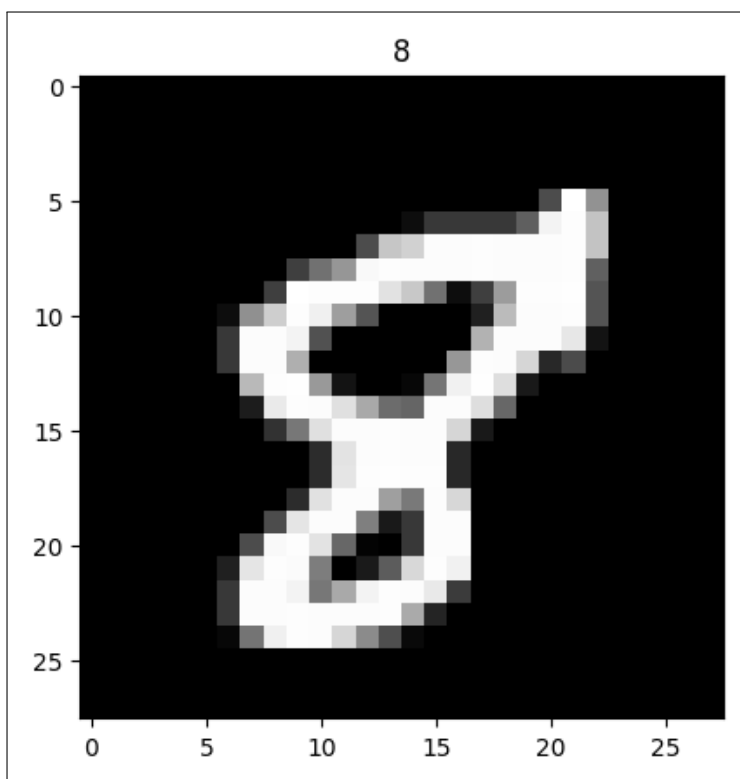
Popis úlohy . . . . .	1
Popis dát a preprocessing . . . . .	1
Model . . . . .	2
Baseline . . . . .	2
Trénovanie . . . . .	4
Testovanie . . . . .	6
Záver a porovnanie s baseline . . . . .	7

## Popis úlohy

Úlohou je naučiť neurónovú sieť identifikovať číslicu (0–9), ktorá sa vyskytuje na obrázku.

## Popis dát a preprocessing

MNIST je veľká databáza ručne písaných čísiel, ktorá sa bežne používa na tréning rôznych systémov spracovania obrazu. Obrázky majú veľkosť 28x28 pixelov. Dataset je voľne prístupný v knižnici Tensorflow pomocou príkazu `tf.keras.datasets.mnist.load_data`. Pozostáva zo 70 000 obrázkov rozdelených na train (60 000) a test (10 000) množiny. Dataset je nevyvážený: pomerné zastúpenie jednotlivých tried vidíme v tabuľke 1. Ukážku jedného obrázku vidíme na obrázku 1.



**Obr. 1:** Ukážka obrázku z MNIST datasetu.

Číslica	Trénovacia množina	Testovacia množina
0	5923	980
1	6742	1135
2	5958	1032
3	6131	1010
4	5842	982
5	5421	892
6	5918	958
7	6265	1028
8	5851	974
9	5949	1009

**Tabuľka 1:** Zastúpenie jednotlivých tried v trénovacej a testovacej množine.

## Model

Nakoľko je MNIST dataset dostatočne veľký nepoužijeme na vyhodnotenie modelu k-fold cross validation. Keďže však dataset nie je vyvážený, použijeme pri trénovaní vypočítané váhy tried, ktoré môžeme vidieť v tabuľke 2.

Číslica	Váha
0	1.0130001688333614
1	0.8899436369029962
2	1.0070493454179255
3	0.978633175664655
4	1.0270455323519343
5	1.1068068622025458
6	1.013856032443393
7	0.9577015163607342
8	1.025465732353444
9	1.0085728693898135

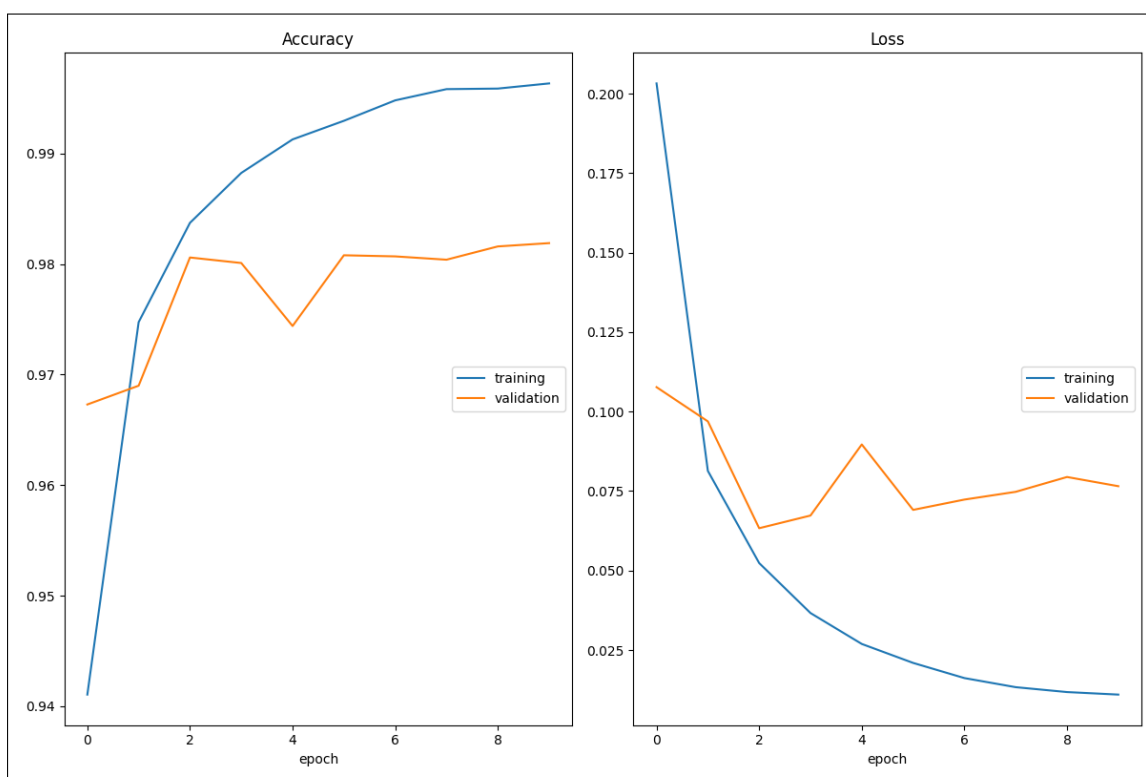
**Tabuľka 2:** Váhy jednotlivých tried pri trénovaní.

## Baseline

Baseline bude jednoduchá neurónová sieť, ktorej architektúru môžeme vidieť v tabuľke 3. Výsledky vidíme v tabuľkách 4, 5. Pribeh trénovania vidíme na obrázku 2. Trénovali sme po dobu 5 epoch a pri batch size 32. Na obrázku 2 vidíme, že model sa veľmi rýchlo (po 2 epochách) pretrénoval.

Name	Type	Shape	# Params	Activation	Regularizer
flatten_1	Flatten	(None, 784)	0	None	None
dense_1	Dense	(None, 512)	401920	relu	None
dense_2	Dense	(None, 10)	5130	softmax	None
Total params:	407050 (1.55 MB)				
Trainable params:	407050 (1.55 MB)				

**Tabuľka 3:** Topológia baseline modelu.



**Obr. 2:** Priebeh tréovania baseline modelu.

	0	1	2	3	4	5	6	7	8	9
0	970	1	1	2	1	1	1	0	2	1
1	0	1128	0	1	0	0	2	0	4	0
2	2	2	1001	6	3	0	2	5	10	1
3	0	0	2	995	0	4	0	3	2	4
4	2	0	3	1	949	1	5	4	2	15
5	2	0	0	8	0	875	3	0	3	1
6	0	1	0	0	1	3	953	0	0	0
7	0	4	4	2	0	0	0	1002	4	12
8	0	1	1	8	0	2	0	3	956	3
9	2	2	0	3	4	3	1	1	3	990

**Tabuľka 4:** Confusion matrix baseline modela.

	precision	recall	f1-score	support
0	0.99	0.99	0.99	980
1	0.99	0.99	0.99	1135
2	0.99	0.97	0.98	1032
3	0.97	0.99	0.98	1010
4	0.99	0.97	0.98	982
5	0.98	0.98	0.98	892
6	0.99	0.99	0.99	958
7	0.98	0.97	0.98	1028
8	0.97	0.98	0.98	974
9	0.96	0.98	0.97	1009
accuracy			0.98	10000
macro avg	0.98	0.98	0.98	10000
weighted avg	0.98	0.98	0.98	10000

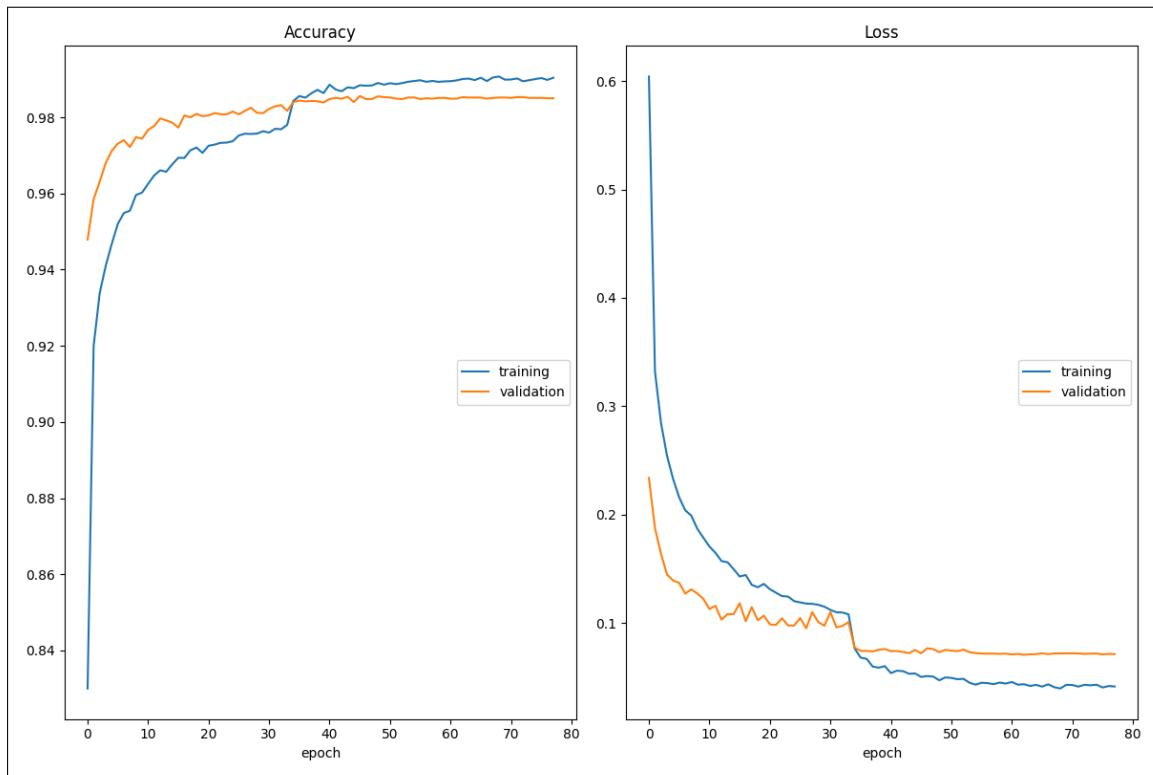
**Tabuľka 5:** Hodnota najbežnejších klasifikačných metrík baseline modela.

## Trénovanie

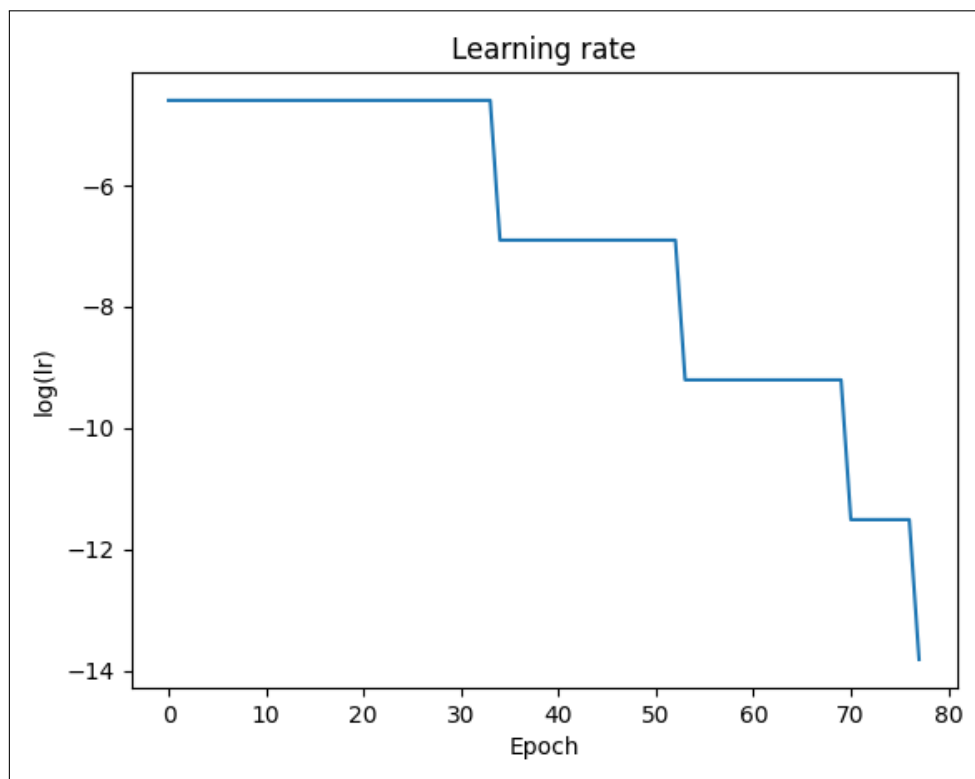
Finálny model môžeme vidieť v tabuľke 6. Priebeh tréovania môžeme vidieť na obrázku 3. Model sme tréovali po dobu 100 epoch (vo výsledku sa však tréoval iba 78 epoch) s počiatočným learning rateom 0,01 a optimalizátorom Adam. Použili sme nasledujúce callbacky: EarlyStopping(patience=15), ReduceLROnPlateau(patience=7). Výsledný priebeh zmeny learning rate vidíme na obrázku 4.

Name	Type	Shape	# Params	Activation	Regularizer
flatten_1	Flatten	(None, 784)	0	None	None
dense_1	Dense	(None, 512)	401920	relu	None
dropout_1	Dropout(0.5)	(None, 512)	0	None	None
dense_2	Dense	(None, 128)	65664	relu	None
dropout_2	Dropout(0.5)	(None, 128)	0	None	None
dense_3	Dense	(None, 32)	4128	relu	None
batch_normalization_1	BatchNormalization	(None, 32)	128	None	None
dropout_3	Dropout(0.1)	(None, 32)	0	None	None
dense_4	Dense	(None, 16)	528	relu	L2=0.01
dense_5	Dense	(None, 10)	170	softmax	None
Total params:		472538 (1.80 MB)			
Trainable params:		472474 (1.80 MB)			
Non-trainable params:		64 (256.00 Byte)			

**Tabuľka 6:** Topológia modelu.



**Obr. 3:** Priebeh tréovania finálneho modelu.



**Obr. 4:** Priebeh hodnoty learning rate počas učenia.

## Testovanie

Confusion matrix predikcie testovacej množiny môžeme vidieť v tabuľke 7. Bežné hodnoty evaluácie modelov nájdeme v tabuľke 8.

	0	1	2	3	4	5	6	7	8	9
0	972	0	0	1	0	1	3	1	2	0
1	0	1125	3	2	0	1	2	0	2	0
2	1	1	1018	1	1	0	1	6	3	0
3	0	0	2	999	0	4	0	3	2	0
4	0	0	3	0	968	0	4	0	0	7
5	2	0	0	3	0	880	2	1	3	1
6	4	2	0	0	1	4	946	0	1	0
7	1	2	8	0	0	1	0	1013	0	3
8	4	0	2	4	3	6	1	2	947	5
9	2	2	0	7	9	1	0	5	1	982

**Tabuľka 7:** Confusion matrix modela.



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

**Tabuľka 8:** Hodnota najbežnejších klasifikačných metrík modelu.

## Záver a porovnanie s baseline

Ako vidíme rozdiel medzi baseline a finálnym modelom nie je významný, avšak napriek tomu finálny model dosahuje lepšie výsledky v precision, recall, a aj f1-score. Rozdiely sú však rádovo v desatinách percent (menovite: 0,4 %, 0,5 %, 0,3 %), čo značí zlepšenie iba o 0,41 %, 0,51 %, 0,31 %. Priemerne sa teda celý model vo všetkých 3 metrikách zlepšil o 0,41 %.

V tabuľke 9 vidíme rozdiely medzi confusion matrix finálneho modelu a baseline modelu. Negatívne hodnoty nesprávnych klasifikácií označujú tie predikcie, v ktorých sa finálny model mýli menej, a naopak pozitívne tie, v ktorých chybuje viac.

Čísla správnych klasifikácií – vyznačená hlavná diagonála – sa interpretujú opačne; pozitívna hodnota rozdielu znamená, že výsledný model určil správne danú číslicu o daný počet klasifikácií častejšie, negatívne číslo zas znamená zhoršenie finálneho modelu.

Za príklad interpretácie tabuľky 9 si môžeme zobrať riadok s číslicou 5. Finálny model si číslicu 5 mýli menej než baseline s číslicami 3 a 6, no viac s číslicou 7. Obecne ju však identifikuje lepšie než baseline model.

	0	1	2	3	4	5	6	7	8	9
0	2	-1	-1	-1	-1	0	2	1	0	-1
1	0	-3	3	1	0	1	0	0	-2	0
2	-1	-1	17	-5	-2	0	-1	1	-7	-1
3	0	0	0	4	0	0	0	0	0	-4
4	-2	0	0	-1	19	-1	-1	-4	-2	-8
5	0	0	0	-5	0	5	-1	1	0	0
6	4	1	0	0	0	1	-7	0	1	0
7	1	-2	4	-2	0	1	0	11	-4	-9
8	4	-1	1	-4	3	4	1	-1	-9	2
9	0	0	0	4	5	-2	-1	4	-2	-8

**Tabuľka 9:** Rozdiel confusion matrix modela a baselinemodela.