

# **VMMT3 Projekt**

## **Cifar10**

# Obsah

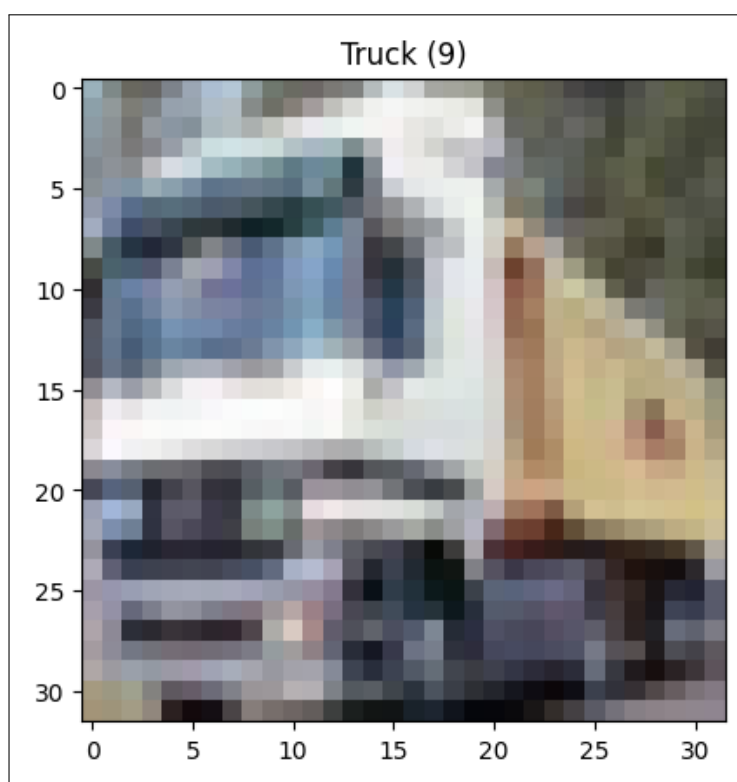
Popis úlohy . . . . .	1
Popis dát a preprocessing . . . . .	1
Model . . . . .	1
Baseline . . . . .	2
Trénovanie . . . . .	3
Testovanie . . . . .	5
Záver a porovnanie s baseline . . . . .	6

## Popis úlohy

Úlohou je naučiť neurónovú sieť identifikovať objekt na fotke.

## Popis dát a preprocessing

Cifar10 je veľká databáza fotografií, ktorá sa bežne používajú na tréovanie rôznych systémov spracovania obrazu. Obrázky majú veľkosť 32x32 pixelov a sú farebné. Dataset je voľne prístupný v knižnici Tensorflow pomocou príkazu `tf.keras.datasets.cifar10.load_data`. Pozostáva zo 60 000 obrázkov rozdelených na train (50 000) a test (10 000) množiny. Dataset je vyvážený, skladá sa z 10 tried po 5 000 vzoriek. Ukážku jedného obrázka aj s popisom vidíme na obrázku 1.



**Obr. 1:** Ukážka obrázku zo Cifar10 datasetu.

## Model

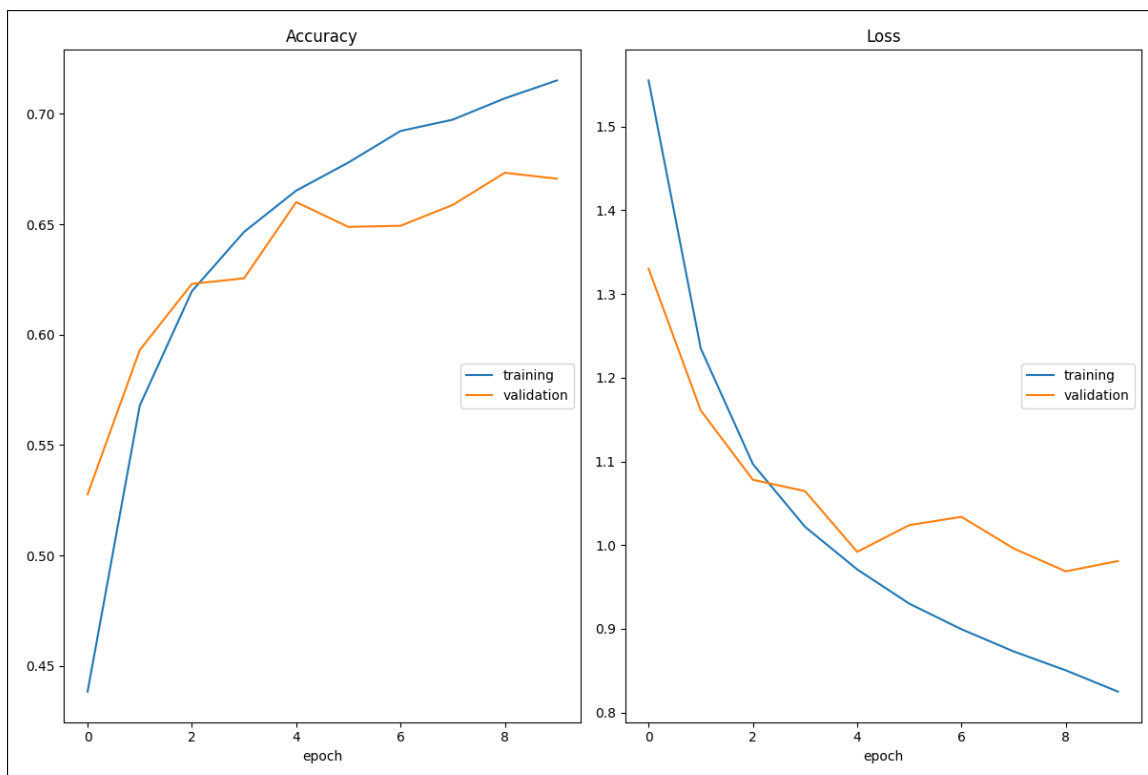
Pre lepšie vyhodnotenie presnosti modelu použijeme pri tréovaní finálneho modelu 5-fold cross validation.

## Baseline

Baseline bude jednoduchá konvolučná neurónová sieť, ktorej architektúru môžeme vidieť v tabuľke 1. Výsledky vidíme v tabuľkách 2, 3. Priebeh tréningovania 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 a spomalil svoju schopnosť zobecňovať.

Name	Type	Shape	#Params	Act.	Reg.	Padding
conv2d	Conv2D	(None, 32, 32, 16)	1216	relu	None	same
max_pooling2d	MaxPooling2D	(None, 16, 16, 16)	0	None	None	–
conv2d_1	Conv2D	(None, 12, 12, 32)	12832	relu	None	valid
max_pooling2d_1	MaxPooling2D	(None, 6, 6, 32)	0	None	None	–
flatten	Flatten	(None, 1152)	0	None	None	–
dense	Dense	(None, 10)	11530	softmax	None	–
Total par.:	25578	(99.91 KB)				
Trainable par.:	25578	(99.91 KB)				
Non-trainable par.:	0	(0.00 Byte)				

**Tabuľka 1:** Topológia baseline modelu.



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

	airplane	automobile	bird	cat	deer	dog	frog	horse	ship	truck
airplane	738	25	82	25	26	19	7	14	78	37
automobile	12	781	5	8	2	4	5	2	27	106
bird	33	6	464	38	54	40	33	31	11	5
cat	23	27	96	497	75	181	78	43	27	21
deer	22	6	134	94	633	50	76	72	7	10
dog	9	8	89	188	55	591	46	81	10	17
frog	13	13	62	69	48	42	734	13	5	14
horse	11	4	39	49	87	57	10	729	2	20
ship	95	39	16	18	17	10	4	5	805	36
truck	44	91	13	14	3	6	7	10	28	734

**Tabuľka 2:** Confusion matrix baseline modela. Zvýraznená hlavná diagonála označuje správne klasifikácie.

	precision	recall	f1-score	support
airplane	0.74	0.70	0.72	1051
automobile	0.78	0.82	0.80	952
bird	0.46	0.65	0.54	715
cat	0.50	0.47	0.48	1068
deer	0.63	0.57	0.60	1104
dog	0.59	0.54	0.56	1094
frog	0.73	0.72	0.73	1013
horse	0.73	0.72	0.73	1008
ship	0.81	0.77	0.79	1045
truck	0.73	0.77	0.75	950
accuracy			0.67	10000
macro avg	0.67	0.67	0.67	10000
weighted avg	0.67	0.67	0.67	10000

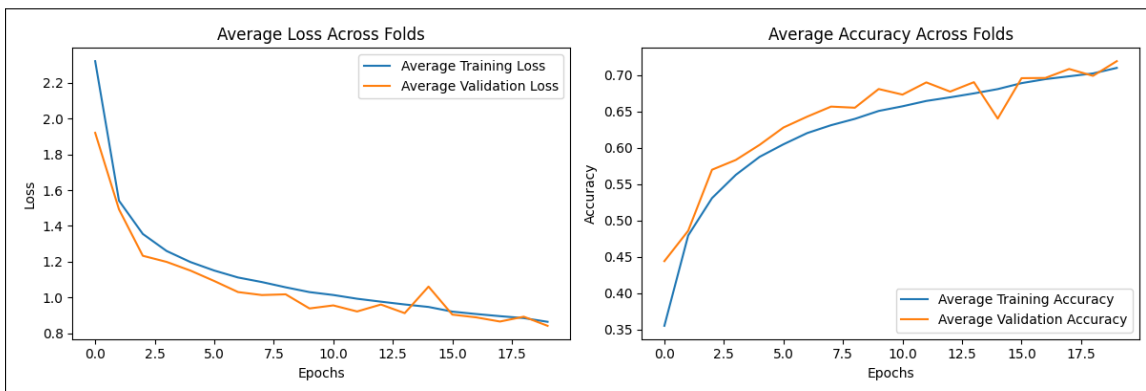
**Tabuľka 3:** Hodnota najbežnejších klasifikačných metrík baseline modela.

## Trénovanie

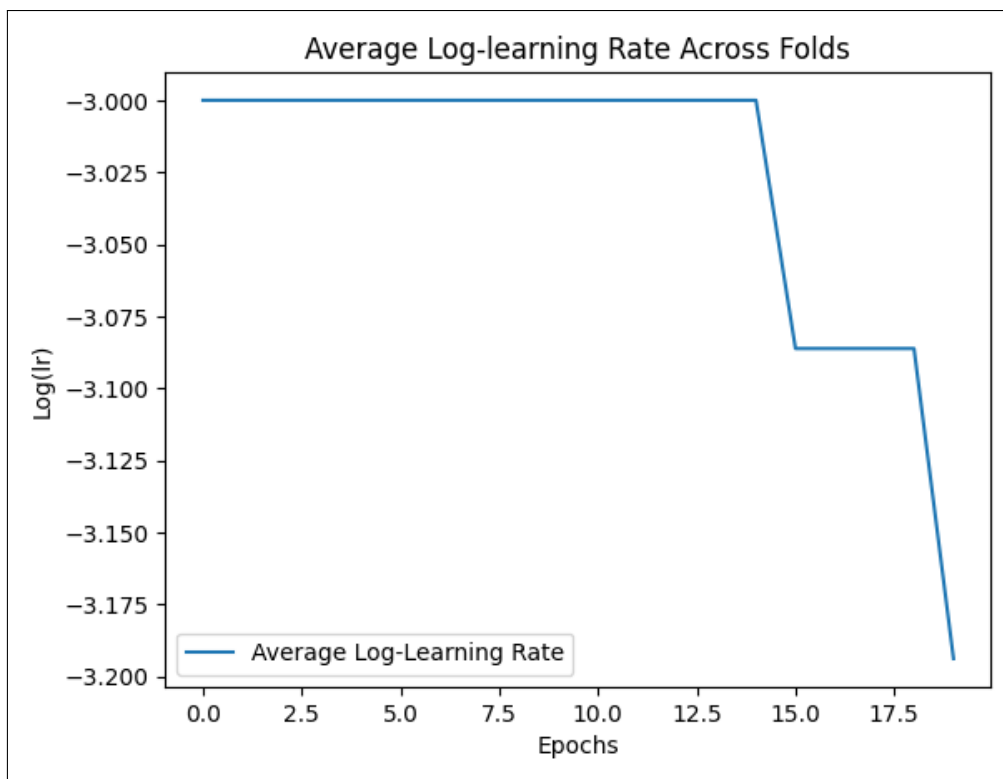
Finálny model môžeme vidieť v tabuľke 4. Priebeh tréovania môžeme vidieť na obrázku 3. Model sme tréovali pomocou 5-fold cross validation (rozdelenia tréovacieho datasetu na validačný a tréovací). Trénovanie trvalo 20 epoch s počiatočným learning rateom 0,001 a optimalizátorom Adam. Použili sme nasledujúce callbacky: ReduceLROnPlateau(patience=5). Výsledný priebeh zmeny learning rate vidíme na obrázku 4.

Name	Type	Shape	#Par.	Act.	Reg.	Padding
conv2d_1	Conv2D	(None, 32, 2432, 32, 32)	2432	relu	None	same
max_pooling2d_1	MaxPooling2D	(None, 16, 0, 16, 32)	0	None	None	–
conv2d_2	Conv2D	(None, 12, 51264, 12, 64)	51264	relu	None	valid
max_pooling2d_2	MaxPooling2D	(None, 6, 0, 6, 64)	0	None	None	–
dropout_1	Dropout(0.5)	(None, 6, 0, 6, 64)	0	None	None	–
conv2d_3	Conv2D	(None, 2, 204928, 2, 128)	204928	relu	None	valid
max_pooling2d_3	MaxPooling2D	(None, 1, 0, 1, 128)	0	None	None	–
batch_normalization	BatchNormalization	(None, 1, 512, 1, 128)	512	None	None	–
flatten_1	Flatten	(None, 128)	0	None	None	–
dropout_2	Dropout(0.5)	(None, 128)	0	None	None	–
dense_1	Dense	(None, 128)	16512	relu	L2=0.01	–
dense_2	Dense	(None, 10)	1290	softmax	None	–
Total par.:	276938	(1.06 MB)				
Trainable par.:	276682	(1.06 MB)				
Non-trainable par.:	256	(1.00 KB)				

**Tabuľka 4:** Topológia modelu.



**Obr. 3:** Priemerný priebeh trénovania finálneho modelu pomocou 5-fold cross-validation.



**Obr. 4:** Priemerný priebeh hodnoty learning rate počas učenia.

## Testovanie

Confusion matrix predikcie testovacej množiny môžeme vidieť v tabuľke 5. Bežné hodnoty evaluácie modelov nájdeme v tabuľke 6.

	airplane	automobile	bird	cat	deer	dog	frog	horse	ship	truck
airplane	724	15	53	8	8	3	4	5	50	21
automobile	13	838	1	1	1	1	0	0	17	45
bird	72	6	583	63	58	46	25	37	21	8
cat	31	11	62	570	49	200	50	44	21	15
deer	37	4	109	74	759	59	31	94	15	10
dog	7	4	66	153	17	620	12	71	5	4
frog	10	23	83	88	58	33	870	7	5	12
horse	8	3	31	23	42	33	4	731	2	6
ship	57	16	8	11	7	1	2	0	841	21
truck	41	80	4	9	1	4	2	11	23	858

**Tabuľka 5:** Confusion matrix modelu. Zvýraznená hlavná diagonála označuje správne klasifikácie.

	precision	recall	f1-score	support
airplane	0.72	0.81	0.77	891
automobile	0.84	0.91	0.87	917
bird	0.58	0.63	0.61	919
cat	0.57	0.54	0.56	1053
deer	0.76	0.64	0.69	1192
dog	0.62	0.65	0.63	959
frog	0.87	0.73	0.79	1189
horse	0.73	0.83	0.78	883
ship	0.84	0.87	0.86	964
truck	0.86	0.83	0.84	1033
accuracy			0.74	10000
macro avg	0.74	0.74	0.74	10000
weighted avg	0.74	0.74	0.74	10000

**Tabuľka 6:** Hodnota najbežnejších klasifikačných metrík modelu.

## Záver a porovnanie s baseline

Ako vidíme rozdiel medzi baseline a finálnym modelom je pomerne významný. Finálny model je vo všetkých metrikách lepší o 10,45 %.

V tabuľke 7 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.



Hodnoty správnych klasifikácií – vyznačená hlavná diagonála – sa interpretujú opačne; pozitívny rozdiel znamená, že výsledný model určil danú triedu o daný počet klasifikácií častejšie než baseline, negatívne číslo zas znamená zhoršenie finálneho modelu.

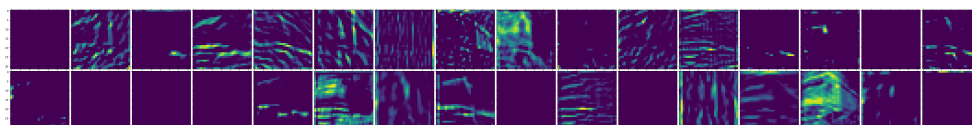
Za príklad interpretácie tabuľky 7 si môžeme zobrať riadok s triedou truck. Finálny model si nákladné auto v porovnaní s baseline modelom mýli menej často so všetkými ostatnými triedami okrem koňa. Obecne teda identifikuje nákladné autá lepšie než baseline model.

Pre lepšiu ilustráciu toho, čo je výstupom jednotlivých konvolučných vrstiev pri spracovaní napríklad obrázku 1, pridávame obrázok 5, na ktorom je vidno postupnú špecifikáciu výstupných filtrov konvolučných vrstiev.

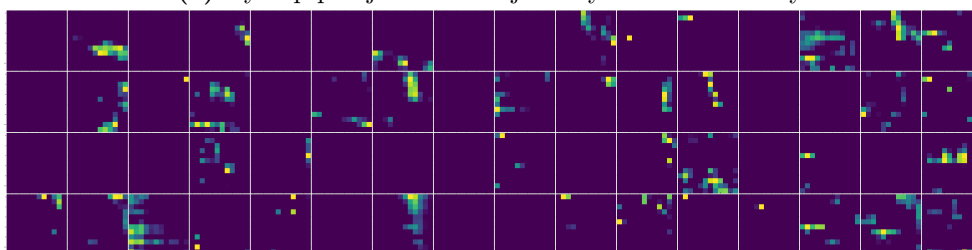
Dá sa predpokladať, že k lepšiemu výsledku (no zároveň k dlhšiemu trénovaniu) by nám pomohla augmentácia dát, ktorá sa dá pomocou knižnice tensorflow jednoducho dosiahnuť či už vytvorením generátora dát pred trénovaním alebo pridaním augmentačných vrstiev na začiatok siete (tieto vrstvy sú aktívne iba pri trénovaní).

	airplane	automobile	bird	cat	deer	dog	frog	horse	ship	truck
airplane	-14	-10	-29	-17	-18	-16	-3	-9	-28	-16
automobile	1	57	-4	-7	-1	-3	-5	-2	-10	-61
bird	39	0	119	25	4	6	-8	6	10	3
cat	8	-16	-34	73	-26	19	-28	1	-6	-6
deer	15	-2	-25	-20	126	9	-45	22	8	0
dog	-2	-4	-23	-35	-38	29	-34	-10	-5	-13
frog	-3	10	21	19	10	-9	136	-6	0	-2
horse	-3	-1	-8	-26	-45	-24	-6	2	0	-14
ship	-38	-23	-8	-7	-10	-9	-2	-5	36	-15
truck	-3	-11	-9	-5	-2	-2	-5	1	-5	124

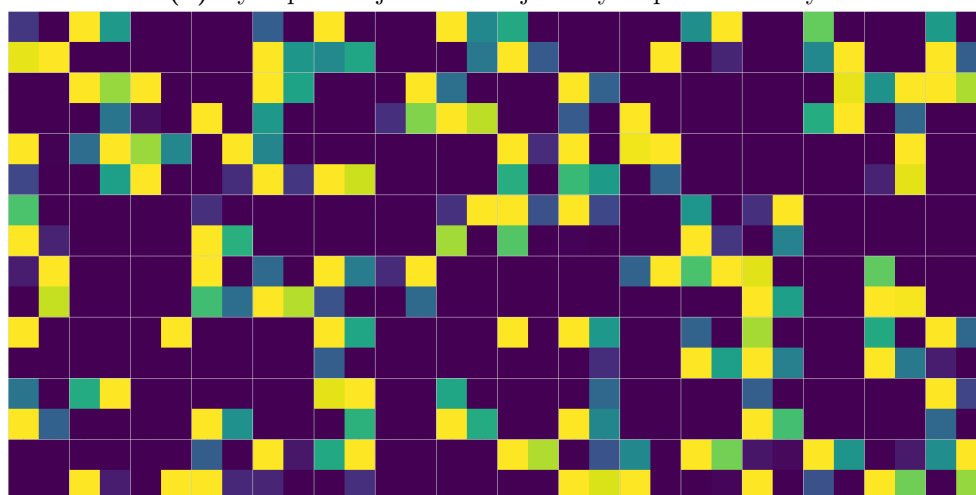
**Tabuľka 7:** Rozdiel confusion matrix modelu a baselinemodelu.



(a) Výstup prvej konvolučnej vrstvy – základné tvary.



(b) Výstup druhej konvolučnej vrstvy – špecifické tvary.



(c) Výstup poslednej konvolučnej vrstvy

**Obr. 5:** Ukážka výstupov jednotlivých konvolučných vrstiev.