

# WSI - ćwiczenie 7.

Sieci neuronowe

grupa 101

Wykonali: Anna Kudzia oraz Adrian Pruszyński.

## 1. Treść Zadania

W ramach szóstego ćwiczenia należy zaimplementować perceptron wielowarstwowy oraz metodę uczącą go przy pomocy algorytmu propagacji wstecznej. Wykorzystując napisaną sieć neuronową należy wytrenować klasyfikator na zbiorze danych MNIST (<http://yann.lecun.com/exdb/mnist/>). Należy pamiętać o podziale zbioru na dane trenujące, walidacyjne i testowe.

## 2. Założenia

Proces uczenia odbywa się przy wykorzystaniu batch'y, których wielkość jest modyfikowalna.

## 3. Prezentacja wyników działania algorytmu

*Tabela 1 Wyniki odstające od pozostałych*

hidden_layers	learning_rate	epochs	learning_time	acc_train	acc_val	acc_test
[25, 10, 10]	0,05	200	328,02	99,85	95,09	95,02
[25, 10, 25]	0,05	200	396,54	99,92	95,33	94,71
[25, 10, 50]	0,05	200	420,35	99,9	95,23	94,9
[25, 25, 10]	0,05	200	501,57	99,78	94,84	94,63
[25, 25, 25]	0,05	200	515,16	99,93	95,12	95,02

Analizując otrzymane wyniki odnaleziono kilka rekordów, których wartość czasu uczenia mocno różni się od pozostałych. Prawdopodobnie jest to spowodowane przerwą w dostawie prądu oraz pracy urządzenia, na którym przeprowadzane były testy, w trybie oszczędzania energii. Wyniki zamieszczone w tabeli 1 nie zostaną uwzględnione w dalszej analiz

Tabela 2 Poprawne wyniki

hidden_layers	learning_rate	epochs	learning_time	acc_train	acc_val	acc_test	$(acc\_test - avg\_acc\_test)^3 / learning\_time$
[10]	0,01	20	16,68	90,63	90,56	90,24	-2,34
[25]	0,01	20	20,14	91,87	91,91	91,35	-0,59
[50]	0,01	20	23,68	92,1	92,25	91,74	-0,29
[10, 10]	0,01	20	18,82	85,46	85,75	85,08	-33,21
[10, 25]	0,01	20	20,62	88,49	88,46	87,92	-9,03
[25, 10]	0,01	20	21,02	87,72	87,83	86,98	-13,99
[25, 25]	0,01	20	21,48	90,39	90,34	89,81	-2,59
[10, 50]	0,01	20	21,28	89,85	90,15	89,57	-3,14
[50, 10]	0,01	20	24,58	89,66	89,56	89,21	-3,51
[50, 50]	0,01	20	27,94	91	91,06	90,62	-0,98
[25, 50]	0,01	20	23,89	90,83	90,91	90,48	-1,31
[50, 25]	0,01	20	26,69	90,47	90,46	89,89	-1,96
[10, 10, 10]	0,01	20	20,91	67,17	66,65	67,37	-866,02
[10, 10, 25]	0,01	20	22,04	73,37	72,31	73,04	-396,05
[10, 10, 50]	0,01	20	24,96	83,88	84,03	83,57	-40,79
[10, 25, 10]	0,01	20	21,98	74,23	74,81	74,08	-339,94
[10, 25, 25]	0,01	20	22,34	81,61	82,07	81,57	-78,51
[10, 25, 50]	0,01	20	23,69	84,49	84,72	84,24	-34,95
[10, 50, 10]	0,01	20	23,62	76,01	75,8	75,88	-236,76
[10, 50, 25]	0,01	20	26,03	82,04	81,97	82,08	-59,19
[10, 50, 50]	0,01	20	27,81	83,34	83,83	82,94	-43,93
[25, 10, 10]	0,01	20	23,09	68,94	69,72	69,11	-638,46
[25, 10, 25]	0,01	20	24	78,48	78,96	78,72	-138,11
[25, 10, 50]	0,01	20	24,95	82,77	83,03	82,43	-56,31
[25, 25, 10]	0,01	20	24,4	72,66	72,64	72,76	-372,54
[25, 25, 25]	0,01	20	25,06	79,73	79,26	79,77	-106,24
[25, 25, 50]	0,01	20	26,97	84,14	84,41	83,9	-34,15
[25, 50, 10]	0,01	20	25,74	79,83	79,87	80,29	-92,23
[25, 50, 25]	0,01	20	27,93	84,49	84,68	83,97	-32,27
[25, 50, 50]	0,01	20	29,85	85,65	85,85	84,84	-22,75

hidden_layers	learning_rate	epochs	learning_time	acc_train	acc_val	acc_test	(acc_test-avg_acc_test)^3/learning_time
[50, 10, 10]	0,01	20	27,67	71,35	71,6	71,49	-392,21
[50, 10, 25]	0,01	20	28,84	73,16	73,98	73,27	-292,64
[50, 10, 50]	0,01	20	28,6	83,32	84	83,41	-37,32
[50, 25, 10]	0,01	20	27,8	76,48	77,14	75,88	-201,16
[50, 25, 25]	0,01	20	30,07	78,69	79,33	78,12	-124,08
[50, 25, 50]	0,01	20	34,25	84,48	84,57	84,15	-24,87
[50, 50, 10]	0,01	20	33,89	81,92	82,73	81,8	-48,85
[50, 50, 25]	0,01	20	31,64	83,63	83,87	83,54	-32,47
[50, 50, 50]	0,01	20	33,38	84,84	84,7	84,59	-22,13
[10]	0,05	20	16,75	93,5	92,8	92,36	-0,12
[25]	0,05	20	18,75	95,48	94,92	94,65	0,06
[50]	0,05	20	21,74	96,28	95,66	95,41	0,26
[10, 10]	0,05	20	20,75	92,93	92,29	91,83	-0,28
[10, 25]	0,05	20	21,27	94,09	93,37	93,15	-0,01
[25, 10]	0,05	20	22,25	95,79	94,86	94,65	0,05
[25, 25]	0,05	20	23,73	95,81	95,1	94,63	0,04
[10, 50]	0,05	20	23,54	93,75	93,01	92,49	-0,06
[50, 10]	0,05	20	26,24	96,41	95,74	95,33	0,19
[50, 50]	0,05	20	28,49	96,7	96,05	95,8	0,36
[25, 50]	0,05	20	23,98	96,15	95,48	95,19	0,16
[50, 25]	0,05	20	26,8	96,63	95,76	95,7	0,33
[10, 10, 10]	0,05	20	21,58	92,02	90,83	90,57	-1,33
[10, 10, 25]	0,05	20	22,64	92,63	91,94	91,15	-0,67
[10, 10, 50]	0,05	20	24,28	93,32	92,31	91,89	-0,22
[10, 25, 10]	0,05	20	22,19	90,36	89,03	89,16	-4,02
[10, 25, 25]	0,05	20	23,64	92,83	91,81	91,54	-0,39
[10, 25, 50]	0,05	20	26,68	93,29	91,96	91,77	-0,24
[10, 50, 10]	0,05	20	25,87	92,85	91,76	91,32	-0,48
[10, 50, 25]	0,05	20	25,94	93,79	93,09	92,49	-0,06
[10, 50, 50]	0,05	20	26,24	93,44	92,56	92,07	-0,14
[25, 10, 10]	0,05	20	29,52	95,66	94,54	93,94	0,00

hidden_layers	learning_rate	epochs	learning_time	acc_train	acc_val	acc_test	$(acc\_test - avg\_acc\_test)^3 / learning\_time$
[25, 10, 25]	0,05	20	27,4	95,09	94,34	93,83	0,00
[25, 10, 50]	0,05	20	28,38	95,52	94,06	93,79	0,00
[25, 25, 10]	0,05	20	24,62	95,74	94,58	94,32	0,01
[25, 25, 25]	0,05	20	27,67	95,98	94,67	94,73	0,05
[25, 25, 50]	0,05	20	26,67	96,05	94,75	94,29	0,01
[25, 50, 10]	0,05	20	27,6	96,07	94,66	94,54	0,03
[25, 50, 25]	0,05	20	26,69	92,68	92,01	91,49	-0,37
[25, 50, 50]	0,05	20	29,13	96,22	95,13	95,05	0,10
[50, 10, 10]	0,05	20	28	96,49	95,07	95,15	0,13
[50, 10, 25]	0,05	20	28,82	96,71	95,68	95,33	0,17
[50, 10, 50]	0,05	20	29,93	96,9	95,78	95,21	0,13
[50, 25, 10]	0,05	20	28,37	96,36	95,2	95,03	0,10
[50, 25, 25]	0,05	20	29,47	96,49	95,44	95,31	0,16
[50, 25, 50]	0,05	20	31,85	96,95	95,98	95,56	0,23
[50, 50, 10]	0,05	20	30,68	96,19	94,83	94,93	0,07
[50, 50, 25]	0,05	20	31,72	96,45	95,45	95,2	0,12
[50, 50, 50]	0,05	20	34,98	96,53	95,53	95,24	0,12
[10]	0,1	20	16,94	93,51	92,38	92,44	-0,10
[25]	0,1	20	19,04	96,68	95,65	95,37	0,28
[50]	0,1	20	22,06	97,56	96,54	96,21	0,78
[10, 10]	0,1	20	18,51	94,28	93,09	92,64	-0,05
[10, 25]	0,1	20	19,8	94,25	92,82	92,75	-0,03
[25, 10]	0,1	20	22,56	97,05	95,58	95,26	0,19
[25, 25]	0,1	20	22,71	97,34	96,05	95,64	0,36
[10, 50]	0,1	20	20,57	94,83	93,57	93,31	0,00
[50, 10]	0,1	20	24,45	97,8	96,07	95,81	0,42
[50, 50]	0,1	20	28,18	98,21	96,94	96,59	0,92
[25, 50]	0,1	20	22,99	97,41	96,1	95,71	0,39
[50, 25]	0,1	20	26,14	98,07	96,49	96,37	0,79
[10, 10, 10]	0,1	20	21,06	93,03	91,6	91,41	-0,52
[10, 10, 25]	0,1	20	22,15	93,7	92,68	92,4	-0,08

hidden_layers	learning_rate	epochs	learning_time	acc_train	acc_val	acc_test	$(acc\_test - avg\_acc\_test)^3 / learning\_time$
[10, 10, 50]	0,1	20	22,74	94,21	92,94	92,8	-0,03
[10, 25, 10]	0,1	20	21,96	94,27	93,09	92,61	-0,05
[10, 25, 25]	0,1	20	23	92,39	91,7	90,82	-0,96
[10, 25, 50]	0,1	20	24,23	94,69	93,39	93,02	-0,01
[10, 50, 10]	0,1	20	23,43	93,81	92,51	91,93	-0,21
[10, 50, 25]	0,1	20	23,95	93,84	92,25	91,92	-0,21
[10, 50, 50]	0,1	20	27,04	94,53	93,06	92,78	-0,02
[25, 10, 10]	0,1	20	23,42	96,4	94,25	94,47	0,03
[25, 10, 25]	0,1	20	24,86	97,29	95,64	95,11	0,13
[25, 10, 50]	0,1	20	24,66	96,8	95,13	94,66	0,04
[25, 25, 10]	0,1	20	23,18	97,22	95,34	95,05	0,12
[25, 25, 25]	0,1	20	23,95	96,88	94,97	94,88	0,08
[25, 25, 50]	0,1	20	25,05	97,12	95,63	95,03	0,11
[25, 50, 10]	0,1	20	25,45	97,26	95,54	94,97	0,09
[25, 50, 25]	0,1	20	25,22	97,25	95,37	95,21	0,16
[25, 50, 50]	0,1	20	28,07	97,22	95,34	95,06	0,10
[50, 10, 10]	0,1	20	27,39	97,99	96,02	95,92	0,44
[50, 10, 25]	0,1	20	26,82	98,39	96,3	96,05	0,53
[50, 10, 50]	0,1	20	27,97	98,44	96,63	96,16	0,58
[50, 25, 10]	0,1	20	27,34	97,71	95,89	95,43	0,21
[50, 25, 25]	0,1	20	28,65	97,81	96,32	95,74	0,33
[50, 25, 50]	0,1	20	29,76	98,29	96,33	96,29	0,63
[50, 50, 10]	0,1	20	29,6	98,09	96,3	96,02	0,46
[50, 50, 25]	0,1	20	30,57	98,32	96,57	96,52	0,79
[50, 50, 50]	0,1	20	33,32	98,46	96,54	96,24	0,53
[10]	0,5	20	16,63	93,98	92,52	91,93	-0,30
[25]	0,5	20	18,12	98,17	96,02	95,76	0,53
[50]	0,5	20	21,57	99,62	97,23	96,78	1,45
[10, 10]	0,5	20	17,89	94,05	92,7	92,05	-0,22
[10, 25]	0,5	20	18,91	94,97	93,69	92,98	-0,01
[25, 10]	0,5	20	20,22	98,22	95,79	95,78	0,49

hidden_layers	learning_rate	epochs	learning_time	acc_train	acc_val	acc_test	$(acc\_test - avg\_acc\_test)^3 / learning\_time$
[25, 25]	0,5	20	21,06	98,51	96,01	95,83	0,51
[10, 50]	0,5	20	19,95	93,74	92,59	91,79	-0,31
[50, 10]	0,5	20	23,51	99,37	96,85	96,77	1,32
[50, 50]	0,5	20	27	99,73	96,89	96,89	1,28
[25, 50]	0,5	20	22,56	97,98	95,98	95,39	0,24
[50, 25]	0,5	20	24,21	99,43	97,03	96,9	1,44
[10, 10, 10]	0,5	20	21,01	94,53	92,86	92,47	-0,07
[10, 10, 25]	0,5	20	23	93,2	91,78	91,3	-0,55
[10, 10, 50]	0,5	20	22,63	94,92	93,7	92,99	-0,01
[10, 25, 10]	0,5	20	25,26	94,55	92,82	92,66	-0,04
[10, 25, 25]	0,5	20	26,42	94,86	93,31	92,75	-0,03
[10, 25, 50]	0,5	20	26,61	95,33	93,56	93,4	0,00
[10, 50, 10]	0,5	20	25,3	93,82	91,49	91,67	-0,30
[10, 50, 25]	0,5	20	28,36	93,36	91,96	91,69	-0,26
[10, 50, 50]	0,5	20	29,05	92,68	91,1	91,32	-0,42
[25, 10, 10]	0,5	20	26,39	97,78	95,43	95,15	0,13
[25, 10, 25]	0,5	20	27,04	97,32	95,38	94,89	0,07
[25, 10, 50]	0,5	20	30,07	98,04	95,64	95,3	0,15
[25, 25, 10]	0,5	20	28,66	97,27	95,3	94,9	0,07
[25, 25, 25]	0,5	20	27,63	98,34	95,74	95,58	0,27
[25, 25, 50]	0,5	20	30,28	98,13	95,69	95,29	0,15
[25, 50, 10]	0,5	20	29,13	98,09	95,62	95,58	0,25
[25, 50, 25]	0,5	20	30,7	97,9	95,52	95,52	0,22
[25, 50, 50]	0,5	20	33,47	98,32	95,74	95,51	0,20
[50, 10, 10]	0,5	20	30,73	98,41	95,97	95,58	0,24
[50, 10, 25]	0,5	20	31,23	99,43	96,4	96,26	0,58
[50, 10, 50]	0,5	20	32,33	99,36	96,73	96,53	0,75
[50, 25, 10]	0,5	20	33,6	99,26	96,78	96,44	0,66
[50, 25, 25]	0,5	20	32,29	99,57	96,85	96,86	1,04
[50, 25, 50]	0,5	20	33,53	99,44	96,77	96,9	1,04
[50, 50, 10]	0,5	20	35,56	99,47	96,67	96,47	0,64

hidden_layers	learning_rate	epochs	learning_time	acc_train	acc_val	acc_test	(acc_test-avg_acc_test)^3/learning_time
[50, 50, 25]	0,5	20	34,47	99,36	96,51	96,65	0,80
[50, 50, 50]	0,5	20	36,43	99,58	96,87	96,63	0,74
[10]	0,01	50	45,1	92,21	91,83	91,63	-0,18
[25]	0,01	50	47,24	93,97	93,83	93,4	0,00
[50]	0,01	50	59,67	94,59	94,4	93,97	0,00
[10, 10]	0,01	50	51,13	91,91	91,6	91,38	-0,22
[10, 25]	0,01	50	53,8	92,62	92,05	91,88	-0,10
[25, 10]	0,01	50	58,48	93,64	93,26	92,77	-0,01
[25, 25]	0,01	50	61,15	94,19	93,73	93,41	0,00
[10, 50]	0,01	50	56,75	92,19	91,32	91,12	-0,28
[50, 10]	0,01	50	69	94,23	93,79	93,35	0,00
[50, 50]	0,01	50	73,94	94,72	94,53	94,23	0,00
[25, 50]	0,01	50	58,43	94,13	93,87	93,49	0,00
[50, 25]	0,01	50	63,2	94,21	94,03	93,59	0,00
[10, 10, 10]	0,01	50	50,98	84,6	84,47	84,09	-17,03
[10, 10, 25]	0,01	50	54,29	88,6	88,61	87,82	-3,61
[10, 10, 50]	0,01	50	57,71	90,57	89,88	90,02	-0,82
[10, 25, 10]	0,01	50	54,65	90,81	90,23	90,12	-0,79
[10, 25, 25]	0,01	50	57,22	90,95	90,57	89,97	-0,86
[10, 25, 50]	0,01	50	60,15	90,93	90,57	89,63	-1,06
[10, 50, 10]	0,01	50	57,88	90,48	90,24	89,78	-0,99
[10, 50, 25]	0,01	50	62,68	91,86	91,78	90,58	-0,45
[10, 50, 50]	0,01	50	66,26	91,69	91,25	90,7	-0,38
[25, 10, 10]	0,01	50	64,04	91,16	91,08	90,43	-0,51
[25, 10, 25]	0,01	50	68,87	93,13	92,58	92,3	-0,03
[25, 10, 50]	0,01	50	72,78	93,63	92,95	92,46	-0,02
[25, 25, 10]	0,01	50	67,28	92,29	91,97	91,17	-0,22
[25, 25, 25]	0,01	50	64,85	93,19	92,73	92,12	-0,05
[25, 25, 50]	0,01	50	68,11	93,2	92,53	92,31	-0,03
[25, 50, 10]	0,01	50	64,12	91,89	91,2	90,73	-0,38
[25, 50, 25]	0,01	50	71,1	93,63	93,05	92,28	-0,03



hidden_layers	learning_rate	epochs	learning_time	acc_train	acc_val	acc_test	(acc_test-avg_acc_test)^3/learning_time
[25, 50, 50]	0,01	50	77,61	93,59	92,81	92,61	-0,01
[50, 10, 10]	0,01	50	71,67	92,82	92,05	92,02	-0,06
[50, 10, 25]	0,01	50	73,9	93,22	92,81	92,59	-0,02
[50, 10, 50]	0,01	50	77,55	94,37	93,89	93,66	0,00
[50, 25, 10]	0,01	50	74,23	93,01	92,72	92,2	-0,04
[50, 25, 25]	0,01	50	77,15	94,06	93,64	93,45	0,00
[50, 25, 50]	0,01	50	82,49	94,02	93,67	93,13	0,00
[50, 50, 10]	0,01	50	82,32	93,22	93,24	92,64	-0,01
[50, 50, 25]	0,01	50	84,53	93,96	93,45	92,99	0,00
[50, 50, 50]	0,01	50	88,7	93,78	93,5	92,89	0,00
[10]	0,05	50	42,99	94,09	92,62	92,47	-0,04
[25]	0,05	50	50,2	97,19	95,85	95,52	0,13
[50]	0,05	50	58,19	98,16	96,88	96,62	0,46
[10, 10]	0,05	50	50,08	95,02	93,27	92,98	-0,01
[10, 25]	0,05	50	52,21	95	93,44	93,36	0,00
[25, 10]	0,05	50	55,5	97,53	95,7	95,61	0,14
[25, 25]	0,05	50	61,57	98,1	96,09	95,92	0,20
[10, 50]	0,05	50	56,35	95,37	93,61	93,55	0,00
[50, 10]	0,05	50	65,3	98,77	96,98	96,72	0,45
[50, 50]	0,05	50	73,47	98,63	96,86	96,64	0,37
[25, 50]	0,05	50	63,23	97,91	96,21	95,96	0,20
[50, 25]	0,05	50	70,82	98,67	96,71	96,66	0,39
[10, 10, 10]	0,05	50	57,15	93,71	92,23	91,71	-0,12
[10, 10, 25]	0,05	50	58,53	95,03	93,27	93,18	0,00
[10, 10, 50]	0,05	50	61,88	95,18	93,29	92,78	-0,01
[10, 25, 10]	0,05	50	64,79	95,34	93,74	93,6	0,00
[10, 25, 25]	0,05	50	66,72	94,86	93,48	92,61	-0,02
[10, 25, 50]	0,05	50	69,32	95,34	93,78	93,37	0,00
[10, 50, 10]	0,05	50	65,03	94,49	92,97	92,39	-0,03
[10, 50, 25]	0,05	50	68,57	94,8	93,26	92,7	-0,01
[10, 50, 50]	0,05	50	74,42	95,39	93,58	93,51	0,00

hidden_layers	learning_rate	epochs	learning_time	acc_train	acc_val	acc_test	(acc_test-avg_acc_test)^3/learning_time
[25, 10, 10]	0,05	50	65,85	97,61	94,84	94,62	0,01
[25, 10, 25]	0,05	50	68,17	97,59	95,22	94,86	0,03
[25, 10, 50]	0,05	50	73,89	98,01	95,33	95,04	0,04
[25, 25, 10]	0,05	50	74,11	97,64	95,31	94,7	0,02
[25, 25, 25]	0,05	50	73,81	97,93	95,7	95,11	0,04
[25, 25, 50]	0,05	50	76,33	98,06	95,43	95,3	0,06
[25, 50, 10]	0,05	50	72,93	97,75	94,8	94,83	0,02
[25, 50, 25]	0,05	50	75,98	98,1	95,67	95,27	0,06
[25, 50, 50]	0,05	50	80,27	98,29	95,93	95,58	0,09
[50, 10, 10]	0,05	50	78,19	98,9	96,48	96,11	0,20
[50, 10, 25]	0,05	50	83,11	98,9	96,34	96,07	0,17
[50, 10, 50]	0,05	50	89,37	99,12	96,62	96,31	0,22
[50, 25, 10]	0,05	50	88,46	99,14	96,32	96,36	0,23
[50, 25, 25]	0,05	50	93,01	99,13	96,77	96,36	0,22
[50, 25, 50]	0,05	50	98,76	99,2	96,53	96,43	0,22
[50, 50, 10]	0,05	50	94,66	99,19	96,54	96,25	0,19
[50, 50, 25]	0,05	50	98,23	99,18	96,48	96,46	0,23
[50, 50, 50]	0,05	50	107,07	99,2	96,54	96,3	0,18
[10]	0,1	50	51,71	94,27	93,24	92,66	-0,02
[25]	0,1	50	58,95	98,07	96,08	95,64	0,14
[50]	0,1	50	70,41	99,18	97,1	96,95	0,52
[10, 10]	0,1	50	57,31	95,34	93,57	93,43	0,00
[10, 25]	0,1	50	60,31	95,42	93,56	93,07	0,00
[25, 10]	0,1	50	65,4	98,52	95,82	95,54	0,11
[25, 25]	0,1	50	71,34	98,59	96,06	95,83	0,15
[10, 50]	0,1	50	65,22	95,81	93,94	93,65	0,00
[50, 10]	0,1	50	81,79	99,76	97,07	96,7	0,35
[50, 50]	0,1	50	90,2	99,79	97,39	97,07	0,45
[25, 50]	0,1	50	75,27	99,01	96,21	95,77	0,13
[50, 25]	0,1	50	83,39	99,71	97,34	97,11	0,51
[10, 10, 10]	0,1	50	67,12	95,03	93,42	92,89	-0,01

hidden_layers	learning_rate	epochs	learning_time	acc_train	acc_val	acc_test	$(acc\_test - avg\_acc\_test)^3 / learning\_time$
[10, 10, 25]	0,1	50	70,8	95,03	92,96	92,46	-0,02
[10, 10, 50]	0,1	50	74,23	95,47	93,13	93,56	0,00
[10, 25, 10]	0,1	50	69,15	94,84	93,17	92,75	-0,01
[10, 25, 25]	0,1	50	73,95	95,13	93,28	92,98	0,00
[10, 25, 50]	0,1	50	78,3	95,38	93,72	93,21	0,00
[10, 50, 10]	0,1	50	75,05	95,37	93,6	93,01	0,00
[10, 50, 25]	0,1	50	78,13	94,93	93,71	92,42	-0,02
[10, 50, 50]	0,1	50	84,08	95,7	93,6	93,53	0,00
[25, 10, 10]	0,1	50	74,43	95,86	93,7	93,51	0,00
[25, 10, 25]	0,1	50	77,05	98,74	95,55	95,57	0,09
[25, 10, 50]	0,1	50	81,28	98,76	95,64	95,37	0,06
[25, 25, 10]	0,1	50	78,79	98,55	95,52	95,38	0,07
[25, 25, 25]	0,1	50	72,01	98,7	95,78	95,59	0,10
[25, 25, 50]	0,1	50	77,42	97,37	94,7	94,77	0,02
[25, 50, 10]	0,1	50	72,43	98,79	95,76	95,78	0,14
[25, 50, 25]	0,1	50	81,52	98,78	95,58	95,33	0,06
[25, 50, 50]	0,1	50	81,96	98,37	95,51	95,16	0,04
[50, 10, 10]	0,1	50	76,49	99,4	96,26	96,2	0,22
[50, 10, 25]	0,1	50	79,7	99,87	96,25	96,45	0,28
[50, 10, 50]	0,1	50	83,13	99,82	96,88	96,44	0,27
[50, 25, 10]	0,1	50	82,42	99,7	96,56	96,44	0,27
[50, 25, 25]	0,1	50	84,9	99,84	96,93	96,48	0,27
[50, 25, 50]	0,1	50	88,62	99,82	96,57	96,63	0,30
[50, 50, 10]	0,1	50	85,46	99,59	96,67	96,35	0,24
[50, 50, 25]	0,1	50	88,55	99,9	96,96	96,58	0,29
[50, 50, 50]	0,1	50	94,7	99,9	96,8	96,91	0,37
[10]	0,5	50	45,71	95,3	93,32	92,87	-0,01
[25]	0,5	50	53	99,35	96,29	95,82	0,20
[50]	0,5	50	65,38	100	97,27	96,89	0,53
[10, 10]	0,5	50	52,39	95,32	92,99	92,45	-0,03
[10, 25]	0,5	50	54,96	95,05	92,25	92,21	-0,05

hidden_layers	learning_rate	epochs	learning_time	acc_train	acc_val	acc_test	$(acc\_test - avg\_acc\_test)^3 / learning\_time$
[25, 10]	0,5	50	59,37	99,4	95,93	95,63	0,13
[25, 25]	0,5	50	61,74	99,66	95,85	95,7	0,14
[10, 50]	0,5	50	58,61	96,27	93,88	93,35	0,00
[50, 10]	0,5	50	69,05	99,99	97,08	96,96	0,53
[50, 50]	0,5	50	82,15	100	97,17	96,98	0,46
[25, 50]	0,5	50	65,98	99,65	96,06	95,78	0,15
[50, 25]	0,5	50	73,47	100	97,3	96,72	0,40
[10, 10, 10]	0,5	50	57,79	94,36	92,18	91,55	-0,16
[10, 10, 25]	0,5	50	61,82	94,47	92,25	91,92	-0,08
[10, 10, 50]	0,5	50	57,56	95,45	93,3	92,57	-0,02
[10, 25, 10]	0,5	50	53,2	95,42	93,12	93,02	0,00
[10, 25, 25]	0,5	50	55,13	95,95	93,4	92,7	-0,01
[10, 25, 50]	0,5	50	58,49	95,98	93,33	93,39	0,00
[10, 50, 10]	0,5	50	56,87	95,43	93,19	92,79	-0,01
[10, 50, 25]	0,5	50	61,65	95,2	92,72	92,85	-0,01
[10, 50, 50]	0,5	50	64,5	95,59	92,92	92,79	-0,01
[25, 10, 10]	0,5	50	56,81	98,46	95,25	95,04	0,05
[25, 10, 25]	0,5	50	59,56	99,11	95,83	95,71	0,15
[25, 10, 50]	0,5	50	63,37	98,79	95,06	95,24	0,07
[25, 25, 10]	0,5	50	59,37	99,33	95,65	95,51	0,11
[25, 25, 25]	0,5	50	61,8	99,24	95,98	95,72	0,15
[25, 25, 50]	0,5	50	65,28	99,38	95,47	95,24	0,06
[25, 50, 10]	0,5	50	63,21	99,16	95,93	95,41	0,09
[25, 50, 25]	0,5	50	65,62	99,37	95,93	95,41	0,09
[25, 50, 50]	0,5	50	71,81	99,2	95,93	95,54	0,10
[50, 10, 10]	0,5	50	66,46	99,94	96,63	96,27	0,28
[50, 10, 25]	0,5	50	69,75	99,98	96,65	96,7	0,41
[50, 10, 50]	0,5	50	72,42	99,99	96,71	96,79	0,44
[50, 25, 10]	0,5	50	70,05	99,97	96,99	96,87	0,49
[50, 25, 25]	0,5	50	71,63	100	97,03	96,66	0,39
[50, 25, 50]	0,5	50	75,21	100	97,39	96,93	0,48

hidden_layers	learning_rate	epochs	learning_time	acc_train	acc_val	acc_test	(acc_test-avg_acc_test)^3/learning_time
[50, 50, 10]	0,5	50	75,24	99,98	97,1	96,81	0,43
[50, 50, 25]	0,5	50	80,44	100	96,87	96,83	0,41
[50, 50, 50]	0,5	50	84,52	100	97,04	96,92	0,42
[10]	0,01	100	83,93	93,44	92,9	92,85	-0,01
[25]	0,01	100	92,3	95,59	95,06	94,77	0,02
[50]	0,01	100	110,61	96,19	95,63	95,27	0,04
[10, 10]	0,01	100	94,67	93,4	92,52	92,26	-0,03
[10, 25]	0,01	100	102,16	93,88	93,03	93,03	0,00
[25, 10]	0,01	100	109,93	96,02	95,02	94,87	0,02
[25, 25]	0,01	100	116,3	95,99	95	94,76	0,01
[10, 50]	0,01	100	122,2	94,25	93,05	93,1	0,00
[50, 10]	0,01	100	130,28	96,66	95,82	95,7	0,07
[50, 50]	0,01	100	150,71	96,68	95,98	95,61	0,05
[25, 50]	0,01	100	128,66	96,15	95,15	95	0,02
[50, 25]	0,01	100	141,79	96,77	96,04	95,71	0,06
[10, 10, 10]	0,01	100	110,78	92,17	90,58	90,52	-0,27
[10, 10, 25]	0,01	100	120,52	92,74	91,64	91,32	-0,10
[10, 10, 50]	0,01	100	125,57	93,36	91,84	92,03	-0,03
[10, 25, 10]	0,01	100	118,29	93,51	92,05	92,21	-0,02
[10, 25, 25]	0,01	100	122,32	93,24	91,98	91,87	-0,04
[10, 25, 50]	0,01	100	126,13	92,79	91,9	91,14	-0,12
[10, 50, 10]	0,01	100	120,09	93,12	91,82	91,48	-0,08
[10, 50, 25]	0,01	100	128,66	93,07	91,74	91,49	-0,08
[10, 50, 50]	0,01	100	135,2	93,91	92,78	92,37	-0,01
[25, 10, 10]	0,01	100	120,72	96,18	94,92	94,39	0,00
[25, 10, 25]	0,01	100	125,17	95,81	94,56	93,78	0,00
[25, 10, 50]	0,01	100	132,99	96,41	95,25	94,92	0,02
[25, 25, 10]	0,01	100	133,01	95,74	94,94	94,36	0,00
[25, 25, 25]	0,01	100	133,5	96,23	94,79	94,64	0,01
[25, 25, 50]	0,01	100	141,68	96,29	94,76	94,62	0,01
[25, 50, 10]	0,01	100	134,92	96,22	94,92	94,53	0,01

hidden_layers	learning_rate	epochs	learning_time	acc_train	acc_val	acc_test	$(acc\_test - avg\_acc\_test)^3 / learning\_time$
[25, 50, 25]	0,01	100	142,14	96,13	94,89	94,56	0,01
[25, 50, 50]	0,01	100	150,74	96,33	95,14	94,91	0,01
[50, 10, 10]	0,01	100	140,97	96,73	95,58	95,02	0,02
[50, 10, 25]	0,01	100	147,58	97,18	95,88	95,52	0,05
[50, 10, 50]	0,01	100	161,43	96,9	95,47	95,21	0,02
[50, 25, 10]	0,01	100	159,03	96,58	95,26	94,95	0,01
[50, 25, 25]	0,01	100	166,41	96,81	95,8	95,29	0,03
[50, 25, 50]	0,01	100	181,89	97,08	96,07	95,54	0,04
[50, 50, 10]	0,01	100	181,3	96,88	95,65	95,46	0,03
[50, 50, 25]	0,01	100	175,19	96,94	95,9	95,45	0,03
[50, 50, 50]	0,01	100	186,9	97,04	95,7	95,62	0,04
[10]	0,05	100	89,48	95,13	93,47	92,8	-0,01
[25]	0,05	100	98,99	98,08	96,03	95,64	0,08
[50]	0,05	100	118,39	99,18	97,08	96,75	0,26
[10, 10]	0,05	100	99,75	95,54	93,55	93,19	0,00
[10, 25]	0,05	100	105,7	95,85	93,59	93,51	0,00
[25, 10]	0,05	100	116,95	99,07	96,12	95,8	0,09
[25, 25]	0,05	100	130,38	98,59	96,01	95,55	0,05
[10, 50]	0,05	100	120,89	95,88	93,65	93,53	0,00
[50, 10]	0,05	100	145,49	99,78	96,93	96,92	0,24
[50, 50]	0,05	100	163,76	99,81	97,29	96,89	0,21
[25, 50]	0,05	100	141,01	99,06	96,38	96,01	0,10
[50, 25]	0,05	100	158,91	99,81	97,22	97,05	0,25
[10, 10, 10]	0,05	100	114,79	95,63	93,62	92,98	0,00
[10, 10, 25]	0,05	100	118,23	95,53	92,96	92,55	-0,01
[10, 10, 50]	0,05	100	125,52	95,35	93,27	92,94	0,00
[10, 25, 10]	0,05	100	120,95	95,54	93,31	92,82	0,00
[10, 25, 25]	0,05	100	127,47	95,85	93,48	93	0,00
[10, 25, 50]	0,05	100	141,2	95,17	93,08	92,64	-0,01
[10, 50, 10]	0,05	100	135,63	95,02	92,87	92,59	-0,01
[10, 50, 25]	0,05	100	144,98	95,71	93,56	93,4	0,00

hidden_layers	learning_rate	epochs	learning_time	acc_train	acc_val	acc_test	(acc_test-avg_acc_test)^3/learning_time
[10, 50, 50]	0,05	100	157,72	95,94	94,07	93,67	0,00
[25, 10, 10]	0,05	100	138,42	99,14	95,22	95,24	0,03
[25, 10, 25]	0,05	100	140,56	98,79	95,29	94,88	0,01
[25, 10, 50]	0,05	100	138,14	98,86	95,17	94,85	0,01
[25, 25, 10]	0,05	100	131,85	99	95,43	95,06	0,02
[25, 25, 25]	0,05	100	136,47	99,3	95,69	95,24	0,03
[25, 25, 50]	0,05	100	142,98	97,77	95,01	94,62	0,01
[25, 50, 10]	0,05	100	142,56	99,16	95,61	95,7	0,06
[25, 50, 25]	0,05	100	152,98	99,33	95,54	95,23	0,03
[25, 50, 50]	0,05	100	165,48	99,31	95,46	94,96	0,01
[50, 10, 10]	0,05	100	157,26	99,91	96,3	96,14	0,10
[50, 10, 25]	0,05	100	165,42	99,97	96,84	96,86	0,20
[50, 10, 50]	0,05	100	172	99,97	96,36	96,51	0,14
[50, 25, 10]	0,05	100	168,68	99,9	96,54	96,6	0,16
[50, 25, 25]	0,05	100	154,86	99,95	96,7	96,36	0,13
[50, 25, 50]	0,05	100	157,64	99,95	96,48	96,48	0,15
[50, 50, 10]	0,05	100	147,37	99,9	96,24	95,91	0,08
[50, 50, 25]	0,05	100	152,64	99,97	96,57	96,47	0,15
[50, 50, 50]	0,05	100	164,28	99,97	96,74	96,37	0,13
[10]	0,1	100	81,81	95,34	93,5	93,05	0,00
[25]	0,1	100	93,36	99,18	95,94	95,63	0,09
[50]	0,1	100	111,35	99,85	97,18	96,88	0,31
[10, 10]	0,1	100	92,35	95,86	93,43	92,87	0,00
[10, 25]	0,1	100	97,65	95,65	93,24	93,12	0,00
[25, 10]	0,1	100	103,17	99,64	95,99	95,64	0,08
[25, 25]	0,1	100	110,89	99,65	96,05	96,03	0,12
[10, 50]	0,1	100	103,43	96,48	94,29	93,55	0,00
[50, 10]	0,1	100	122,67	99,97	96,98	96,71	0,24
[50, 50]	0,1	100	139,39	99,99	97,13	96,92	0,26
[25, 50]	0,1	100	114,77	99,72	95,8	95,54	0,06
[50, 25]	0,1	100	129,54	100	96,99	96,9	0,27

hidden_layers	learning_rate	epochs	learning_time	acc_train	acc_val	acc_test	$(acc\_test - avg\_acc\_test)^3 / learning\_time$
[10, 10, 10]	0,1	100	104,6	94,64	92,62	92,11	-0,03
[10, 10, 25]	0,1	100	107,41	95,96	93,49	93,13	0,00
[10, 10, 50]	0,1	100	113,57	96,62	93,67	93,48	0,00
[10, 25, 10]	0,1	100	106,75	96,18	93,5	93,3	0,00
[10, 25, 25]	0,1	100	112,46	95,6	93,02	92,6	-0,01
[10, 25, 50]	0,1	100	123,06	95,86	93,5	92,83	0,00
[10, 50, 10]	0,1	100	117,3	96,19	93,99	93,65	0,00
[10, 50, 25]	0,1	100	124,32	96,18	93,74	93,52	0,00
[10, 50, 50]	0,1	100	132,78	95,9	93,44	93,1	0,00
[25, 10, 10]	0,1	100	118,17	99,58	95,33	95,29	0,04
[25, 10, 25]	0,1	100	122	99,55	94,88	95,12	0,03
[25, 10, 50]	0,1	100	128,18	99,63	95,37	95,04	0,02
[25, 25, 10]	0,1	100	122,78	99,49	95,31	95,02	0,02
[25, 25, 25]	0,1	100	127,39	99,79	95,77	95,22	0,03
[25, 25, 50]	0,1	100	136,71	99,77	95,6	95,14	0,03
[25, 50, 10]	0,1	100	129,53	99,41	95,41	95,12	0,03
[25, 50, 25]	0,1	100	135,63	99,62	95,96	95,56	0,05
[25, 50, 50]	0,1	100	146,04	99,77	95,39	95,53	0,05
[50, 10, 10]	0,1	100	136,99	100	96,51	96,42	0,16
[50, 10, 25]	0,1	100	144,22	99,99	96,53	96,48	0,16
[50, 10, 50]	0,1	100	148,12	100	96,69	96,32	0,13
[50, 25, 10]	0,1	100	142,53	99,98	96,43	96,34	0,14
[50, 25, 25]	0,1	100	149,17	100	96,64	96,38	0,14
[50, 25, 50]	0,1	100	154,5	100	96,78	96,77	0,20
[50, 50, 10]	0,1	100	153,84	99,98	96,31	95,89	0,08
[50, 50, 25]	0,1	100	158,14	100	96,68	96,4	0,13
[50, 50, 50]	0,1	100	170,39	100	96,63	96,5	0,14
[10]	0,5	100	81,42	95,64	93,11	92,55	-0,02
[25]	0,5	100	96,45	99,92	95,77	95,26	0,04
[50]	0,5	100	114,26	100	97,17	96,93	0,31
[10, 10]	0,5	100	93,56	94,64	91,19	90,97	-0,20



hidden_layers	learning_rate	epochs	learning_time	acc_train	acc_val	acc_test	$(acc\_test - avg\_acc\_test)^3 / learning\_time$
[10, 25]	0,5	100	97,83	96,45	93,62	92,83	-0,01
[25, 10]	0,5	100	105,3	99,41	95,14	94,97	0,02
[25, 25]	0,5	100	110,71	99,97	95,5	95,41	0,05
[10, 50]	0,5	100	105,46	96,56	93,38	92,73	-0,01
[50, 10]	0,5	100	124,75	100	96,76	96,39	0,17
[50, 50]	0,5	100	140,51	100	97,33	97,22	0,33
[25, 50]	0,5	100	116,39	99,99	95,87	95,72	0,08
[50, 25]	0,5	100	132,1	100	97,24	97,13	0,32
[10, 10, 10]	0,5	100	104,34	95,51	93,05	92,2	-0,03
[10, 10, 25]	0,5	100	110,28	95,84	92,71	92,41	-0,02
[10, 10, 50]	0,5	100	116,24	95,96	93,26	92,54	-0,01
[10, 25, 10]	0,5	100	108,7	94,94	92,49	91,9	-0,05
[10, 25, 25]	0,5	100	116,02	96,76	93,68	92,73	-0,01
[10, 25, 50]	0,5	100	121,3	96,2	93,21	93,02	0,00
[10, 50, 10]	0,5	100	116,03	96,99	93,89	93,57	0,00
[10, 50, 25]	0,5	100	121,87	96,69	93,82	93,13	0,00
[10, 50, 50]	0,5	100	132,79	96,51	93,34	92,69	-0,01
[25, 10, 10]	0,5	100	118,01	98,71	94,32	94,33	0,00
[25, 10, 25]	0,5	100	120,96	98,58	94,9	94,5	0,01
[25, 10, 50]	0,5	100	127,42	98,86	95,18	94,68	0,01
[25, 25, 10]	0,5	100	121,74	99,73	95,64	95,24	0,03
[25, 25, 25]	0,5	100	126,68	99,67	95,37	95,59	0,06
[25, 25, 50]	0,5	100	134,81	99,49	95,64	95,22	0,03
[25, 50, 10]	0,5	100	134,42	99,16	95	94,75	0,01
[25, 50, 25]	0,5	100	135,85	99,91	95,81	95,45	0,04
[25, 50, 50]	0,5	100	143,62	99,93	95,9	95,52	0,05
[50, 10, 10]	0,5	100	135,09	99,99	96,59	96,46	0,17
[50, 10, 25]	0,5	100	144,39	100	96,6	96,37	0,14
[50, 10, 50]	0,5	100	146,72	100	96,88	96,79	0,22
[50, 25, 10]	0,5	100	141,93	100	96,71	96,62	0,19
[50, 25, 25]	0,5	100	146,05	100	96,99	96,99	0,26

hidden_layers	learning_rate	epochs	learning_time	acc_train	acc_val	acc_test	(acc_test-avg_acc_test)^3/learning_time
[50, 25, 50]	0,5	100	154,19	100	97,15	97,01	0,25
[50, 50, 10]	0,5	100	151,74	99,99	96,85	96,74	0,20
[50, 50, 25]	0,5	100	157,66	100	96,99	96,66	0,18
[50, 50, 50]	0,5	100	166,49	100	97,23	97,06	0,24
[10]	0,01	200	164,05	94,36	93,33	93,09	0,00
[25]	0,01	200	187,97	96,75	95,7	95,25	0,02
[50]	0,01	200	226,85	97,7	96,44	96,35	0,09
[10, 10]	0,01	200	187,45	94,92	93,1	93,26	0,00
[10, 25]	0,01	200	219,44	95,07	93,77	93,48	0,00
[25, 10]	0,01	200	213,44	97,47	95,84	95,52	0,03
[25, 25]	0,01	200	222,97	97,77	96,21	95,92	0,05
[10, 50]	0,01	200	209,1	95,77	94,32	93,7	0,00
[50, 10]	0,01	200	248,62	98,49	97,16	96,76	0,12
[50, 50]	0,01	200	284,53	98,42	96,96	96,54	0,09
[25, 50]	0,01	200	234,72	97,5	96,02	95,95	0,05
[50, 25]	0,01	200	265,5	98,45	96,75	96,66	0,10
[10, 10, 10]	0,01	200	208,47	94,48	92,54	92,17	-0,01
[10, 10, 25]	0,01	200	219,41	94,58	92,67	92,2	-0,01
[10, 10, 50]	0,01	200	234,42	95,19	93,37	93,02	0,00
[10, 25, 10]	0,01	200	218,91	94,44	92,33	92,14	-0,02
[10, 25, 25]	0,01	200	231,81	95,23	93,19	93,08	0,00
[10, 25, 50]	0,01	200	242,48	95,17	92,95	92,76	0,00
[10, 50, 10]	0,01	200	230,05	95,06	93,56	92,87	0,00
[10, 50, 25]	0,01	200	242,37	95,08	93,58	93,13	0,00
[10, 50, 50]	0,01	200	263,87	94,73	93,18	93,01	0,00
[25, 10, 10]	0,01	200	235,27	97,87	95,08	94,66	0,00
[25, 10, 25]	0,01	200	243,25	97,61	95,05	94,84	0,01
[25, 10, 50]	0,01	200	258,23	98,35	96,02	95,59	0,03
[25, 25, 10]	0,01	200	247,18	97,57	95,43	94,65	0,00
[25, 25, 25]	0,01	200	253,97	98,01	95,39	95,23	0,02
[25, 25, 50]	0,01	200	273,22	98,09	95,61	95,27	0,02

hidden_layers	learning_rate	epochs	learning_time	acc_train	acc_val	acc_test	(acc_test-avg_acc_test)^3/learning_time
[25, 50, 10]	0,01	200	257,99	98,19	95,9	95,6	0,03
[25, 50, 25]	0,01	200	271,75	97,91	95,34	95,38	0,02
[25, 50, 50]	0,01	200	290,17	97,99	95,54	95,13	0,01
[50, 10, 10]	0,01	200	273,91	98,82	96,08	95,99	0,05
[50, 10, 25]	0,01	200	282,07	99	96,49	96,01	0,05
[50, 10, 50]	0,01	200	295,41	99,22	96,66	96,25	0,06
[50, 25, 10]	0,01	200	284,36	98,97	96,61	96,32	0,07
[50, 25, 25]	0,01	200	297,66	99,09	96,39	96,26	0,06
[50, 25, 50]	0,01	200	307,47	99,08	96,51	96,12	0,05
[50, 50, 10]	0,01	200	306,27	99,08	96,55	96,12	0,05
[50, 50, 25]	0,01	200	317,59	99,02	96,47	96,49	0,07
[50, 50, 50]	0,01	200	339,4	98,98	96,49	96,14	0,05
[10]	0,05	200	173,51	96,15	94,02	93,94	0,00
[25]	0,05	200	185,74	99,1	96,28	95,92	0,06
[50]	0,05	200	224,95	99,88	97,25	96,98	0,17
[10, 10]	0,05	200	187,07	96,11	93,16	93	0,00
[10, 25]	0,05	200	196,87	95,26	92,84	92,43	-0,01
[25, 10]	0,05	200	215,39	99,81	95,82	95,17	0,02
[25, 25]	0,05	200	222,36	99,78	96,24	95,61	0,03
[10, 50]	0,05	200	209,2	96,45	93,97	93,7	0,00
[50, 10]	0,05	200	250,21	99,99	96,97	96,91	0,14
[50, 50]	0,05	200	281,1	100	97,14	96,99	0,13
[25, 50]	0,05	200	237,9	99,85	95,91	95,76	0,04
[50, 25]	0,05	200	259,14	100	97,13	96,75	0,12
[10, 10, 10]	0,05	200	209,28	95,29	93,1	92,01	-0,02
[10, 10, 25]	0,05	200	223,56	95,84	93,14	92,57	-0,01
[10, 10, 50]	0,05	200	231,91	96,44	93,36	93,46	0,00
[10, 25, 10]	0,05	200	221,72	96,65	93,29	93,14	0,00
[10, 25, 25]	0,05	200	230,34	96,93	93,75	93,28	0,00
[10, 25, 50]	0,05	200	242,54	96,44	93,07	93,45	0,00
[10, 50, 10]	0,05	200	232,06	97,18	94,14	93,49	0,00

hidden_layers	learning_rate	epochs	learning_time	acc_train	acc_val	acc_test	$(acc\_test - avg\_acc\_test)^3 / learning\_time$
[10, 50, 25]	0,05	200	243,38	96,58	93,77	93,06	0,00
[10, 50, 50]	0,05	200	268,16	96,61	93,98	93,1	0,00
[25, 25, 50]	0,05	200	273,85	99,96	95,21	94,72	0,00
[25, 50, 10]	0,05	200	280,5	99,81	95,14	95,04	0,01
[25, 50, 25]	0,05	200	288,9	99,95	95,28	95,19	0,01
[25, 50, 50]	0,05	200	310,22	99,95	95,29	95,04	0,01
[50, 10, 10]	0,05	200	288,89	99,99	96,45	96,27	0,06
[50, 10, 25]	0,05	200	308,55	100	96,79	96,55	0,08
[50, 10, 50]	0,05	200	310,07	100	96,68	96	0,04
[50, 25, 10]	0,05	200	298,56	99,99	96,68	96,29	0,06
[50, 25, 25]	0,05	200	315,05	100	96,62	96,35	0,06
[50, 25, 50]	0,05	200	329,41	100	96,71	96,75	0,09
[50, 50, 10]	0,05	200	330,85	99,99	96,54	96,18	0,05
[50, 50, 25]	0,05	200	335,61	100	96,73	96,63	0,08
[50, 50, 50]	0,05	200	373,82	100	96,78	96,67	0,08
[10]	0,1	200	175,12	95,35	92,59	92,45	-0,01
[25]	0,1	200	202,53	99,74	95,9	95,63	0,04
[50]	0,1	200	239,14	99,99	97,14	96,98	0,16
[10, 10]	0,1	200	195,85	96,61	93,06	93,17	0,00
[10, 25]	0,1	200	210,53	96,96	94,07	93,96	0,00
[25, 10]	0,1	200	237,15	99,96	95,73	95,48	0,03
[25, 25]	0,1	200	266,84	99,98	95,7	95,41	0,02
[10, 50]	0,1	200	248,17	97,37	94,18	93,94	0,00
[50, 10]	0,1	200	283,41	100	96,99	96,78	0,11
[50, 50]	0,1	200	325,94	100	97,21	96,93	0,11
[25, 50]	0,1	200	277,22	99,99	96,35	95,92	0,04
[50, 25]	0,1	200	281,64	100	97,12	96,72	0,10
[10, 10, 10]	0,1	200	224,2	96,65	93,54	93,1	0,00
[10, 10, 25]	0,1	200	237,17	96,63	93,3	92,78	0,00
[10, 10, 50]	0,1	200	247,54	96,4	93,04	92,48	0,00

Pomiary dla 200 epok oraz współczynnika uczenia się 0.1 i 0.5 zostały zatrzymane ze względu na występujące już we wcześniejszych pomiarach przeuczenie się sieci. Wyniki w tabeli zostały zaprezentowane poprzez pokolorowanie komórek w zależności od wartości. Od koloru zielonego- najlepsze wyniki do czerwonego- najgorsze. Ponadto w celu łatwiejszego wyszukania najlepszych rozwiązań dodana została dodatkowa kolumna, której wartości wyliczane są zgodnie ze wzorem:

$$Column8 = \frac{(acc_{test} - avg\_acc_{test})^3}{learning\_time}$$

Względem tej kolumny zostały wyznaczone 3 najlepsze wyniki oznaczone poprzez błękitne pokolorowanie komórki w kolumnie ósmej.

Tabela 3 Najlepsze wyniki

hidden_layers	learning_rate	epochs	learning_time	acc_train	acc_val	acc_test	(acc_test-avg_acc_test)^3/learning_time
[50]	0,5	20	21,57	99,62	97,23	96,78	1,45
[50, 10]	0,5	20	23,51	99,37	96,85	96,77	1,32
[50, 25]	0,5	20	24,21	99,43	97,03	96,9	1,44

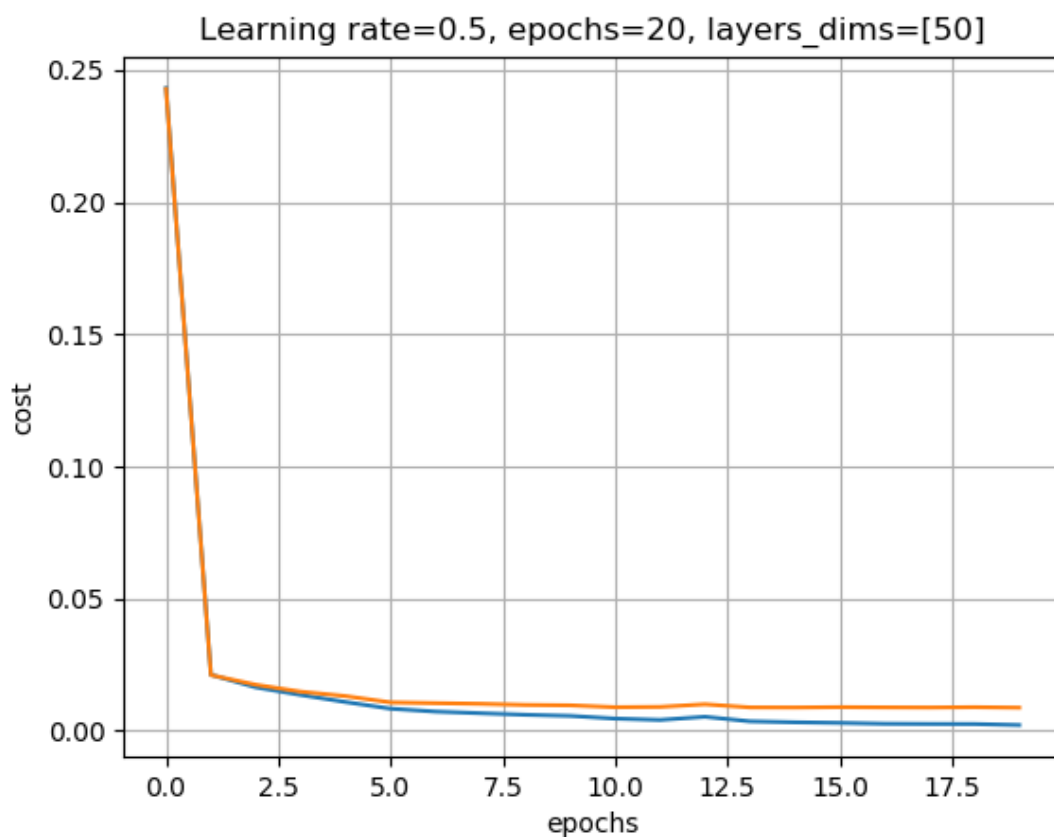
Jak możemy zaobserwować w tabeli 3 najlepsze wyniki obserwujemy dla wariantu 20 epok oraz współczynnik uczenia równy 0,5. Warto również zauważyć, że w każdym z wyników pierwsza warstwa ukryta ma rozmiar 50.

## 4. Analiza działania algorytmu

Rysunek 1 Wykres dla najlepszego wariantu

Zbiór walidacyjny - pomarańczowy.

Zbiór uczący – niebieski

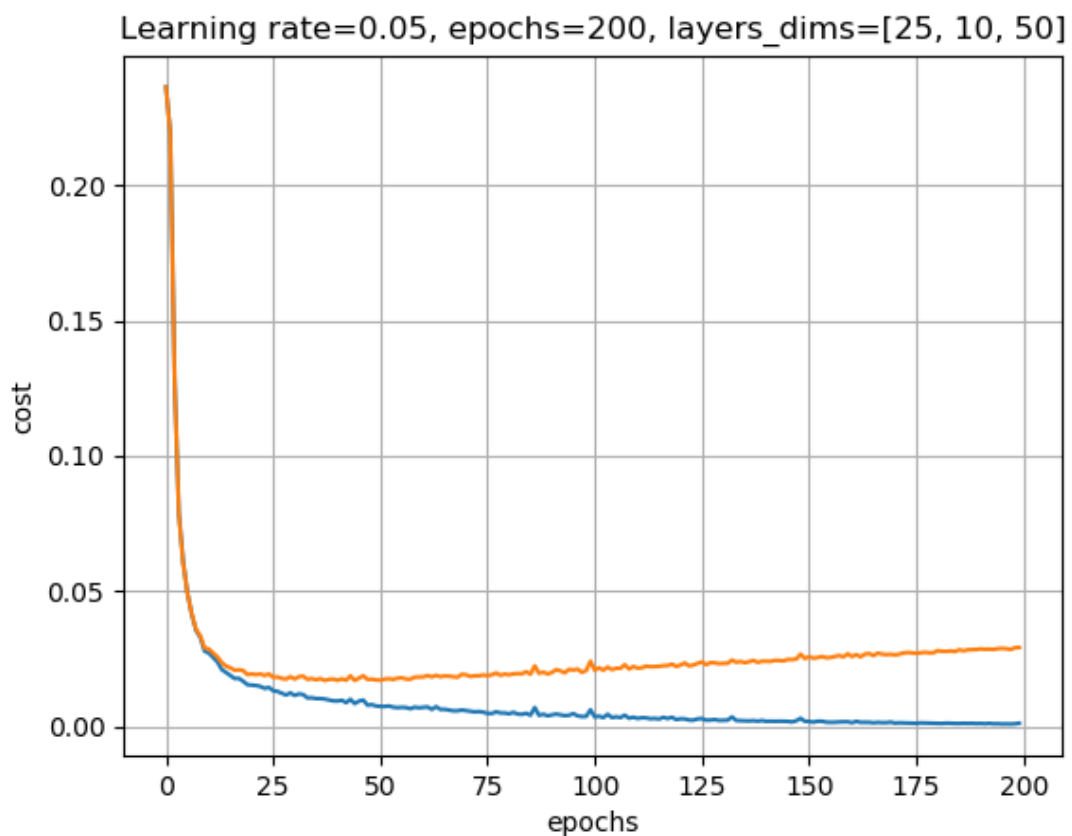


Jak widać na wykresie rysunek 1. Sieć bardzo szybko poprawiła wartość cost, następnie poprawa nie była już tak szybka. Powodem tak szybkiej poprawy w początkowych epokach jest wysoki współczynnik uczenia. W tym przypadku sytuację mógłby poprawić optymalizator np. Adam, dzięki zmniejszaniu wartości współczynnika uczenia w kolejnych epokach.

Rysunek 2 Wykres pokazujący efekt przeuczenia sieci

Zbiór walidacyjny - pomarańczowy.

Zbiór uczący – niebieski



Przyglądając się zmianom wartości cost dla zbioru walidacyjnego łatwo zauważyć, że powyżej pewnej epoki wartość zaczyna wzrastać, pomimo że wartość dla zbioru trenującego maleje. Jest to wyraźny syndrom przeuczenia się sieci przez co sieć dopasowuje się do zbioru trenującego tracąc zdolność do uogólniania.

## 5. Wnioski i podsumowanie

Odpowiednie zaprojektowanie sieci oraz dobór parametrów są kluczowe dla jej efektywności.

Rozmiar sieci powinien być dopasowany do poziomu złożoności zadania. Zbyt mała sieć nie będzie w stanie realizować zadania z zadowalającymi rezultatami. Natomiast zbyt duża sieć będzie wymagać długiego uczenia.

Ilość epok ma ogromne znaczenie w przypadku uczenia sieci, jak można zaobserwować na rysunku 2 zbyt duża ilość epok wydłuża czas uczenia oraz może prowadzić do przeuczenia sieci, natomiast zbyt mała prowadzi do nieoduczenia sieci.

Dobrze dobrany współczynnik uczenia pozwala zaoszczędzić czas wymagany do wyuczenia sieci, ponieważ zbyt mały współczynnik uczenia prowadzi do wzrostu ilości epok wymaganych, aby wyuczyć sieć do zadowalającego poziomu realizacji zadania. Zbyt duża może prowadzić do oscylacji wartości kosztu, w niektórych przypadkach mogą powodować ciągle „przeskakiwanie” wartości optymalnej.