

Neurónové siete - Back Propagation

Juraj Onderik

7.4.2004

Riešil som problém spätného učenia (back propagation) pre sadu D2 (nasob.txt). Vstup som reprezentoval dvoma 3-bitovými číslami a výstup jedným 6-bitovým. Teda vstupná aj výstupná vrstva ma 6 neurónov. Vstupné údaje som náhodne rozdelil v pomere 4:1 na tréningové a testovacie dáta. Akivačná funkcia je sigmoida (vhodná „strmost“ je 3) (**Graf3**)

Optimálnu konfiguráciu siete som hľadal nasledovným spôsobom (sú to priemerné hodnoty z 5 simulácií):

- optimálny počet neurónov v skrytej vrstve (**Tab1**).
- optimálny koeficient momentov - μ (**Tab2**).
- optimálna rýchlosť učenia - λ (**Tab3**).

#	tréningové dáta				testovacie dáta			
	chyba	dobre	neviem	zle	chyba	dobre	neviem	zle
3	13.669289	68.69	26.60	4.71	8.635750	57.44	25.90	16.67
6	7.368804	81.44	16.08	2.48	9.523753	61.28	20.00	18.72
9	2.482514	98.24	0.20	1.57	8.258320	64.87	17.44	17.69
12	1.651638	98.89	0.07	1.05	8.880857	60.51	22.05	17.44
15	1.237465	99.22	0.00	0.78	7.540623	67.44	17.18	15.38
18	0.933009	99.41	0.00	0.59	5.784028	73.85	15.38	10.77
21	1.817587	98.82	0.00	1.18	11.031020	53.33	23.33	23.33

Tab1 - Vývoj učenia vzhľadom na veľkosť skrytej vrstvy ($\lambda=0.1$, $\mu=0.5$, $g=0.1$, epocha=1000)

μ	tréningové dáta				testovacie dáta			
	chyba	dobre	neviem	zle	chyba	dobre	neviem	zle
0.00	0.344418	99.80	0.00	0.20	4.423300	75.90	16.15	7.95
0.20	0.629699	99.61	0.00	0.39	4.101493	78.21	14.10	7.69
0.40	0.421341	99.74	0.00	0.26	4.670143	77.69	12.56	9.74
0.60	1.715747	98.89	0.00	1.11	4.321548	79.74	13.08	7.18
0.80	11.006427	92.81	0.00	7.19	4.957182	82.82	5.38	11.79

Tab2 - Vývoj učenia vzhľadom na μ . ($\lambda=0.1$, $g=0.1$, epocha=1000, skrytá vrstva=19)

λ	tréningové dáta				testovacie dáta			
	chyba	dobre	neviem	zle	chyba	dobre	neviem	zle
0.01	2.234210	70.65	28.95	0.39	3.933172	66.92	27.69	5.38
0.1	0.541675	99.67	0.00	0.33	3.938542	79.23	13.59	7.18
0.2	0.721850	99.54	0.00	0.46	4.673300	77.44	13.85	8.72
0.3	0.612563	99.61	0.00	0.39	5.167386	78.72	12.05	9.23
0.4	0.512344	99.67	0.00	0.33	5.547425	74.62	14.62	10.77
0.5	1.108147	99.28	0.00	0.72	4.979620	80.26	8.97	10.77
0.6	0.807635	99.48	0.00	0.52	5.588581	75.90	13.33	10.77

Tab3 - Vývoj učenia vzhľadom na λ . ($\mu=0$, $g=0.1$, epocha=1000, skrytá vrstva=19)

Najlepšie výsledky pre trénovanie aj testovanie som dosiahol s 19 neurónmi v skrytej vrstve a $\mu=0$ (teda bez použitia momentov) a $\lambda = 0.1$. Ďalšie skryté vrstvy nepreukázali žiadne lepšie výsledky. Je zaujímavé, že keď som učil jednoduchú XOR sieť s dvoma skrytými neurónmi, momentové členy výrazne pomohli (optimálne $\mu=0.3$). Metóda klesajúcich váh mi nefungovala, respektíve učenie nekonvergovalo.

Prehľad učenia siete s optimálnou konfiguráciou uvádzam v **(Tab4)** a **(Tab5)** pre $g=0.1$ a $g=0.2$. Správne, nesprávne a nejasne klasifikované výstupy siete uvádzam v percentuálnej hodnote z celkového počtu bitov výstupov pre danú epochu. Teda ak sú napríklad 3 bity v jednom výsledku nesprávne považujem to za 3 chyby z celkového počtu bitov a nie len ako jeden nesprávny výsledok.

pokús	trénovacie dáta				testovacie dáta			
	Chyba	dobré	neviem	zle	chyba	dobré	neviem	zle
1	0.536673	99.67	0.00	0.33	3.695443	87.18	5.13	7.69
2	0.542664	99.67	0.00	0.33	3.080270	75.64	17.95	6.41
3	0.540438	99.67	0.00	0.33	4.167558	79.49	12.82	7.69
4	0.545650	99.67	0.00	0.33	3.796366	79.49	14.10	6.41
5	0.542953	99.67	0.00	0.33	4.953072	74.36	17.95	7.69

Tab4 - Prehľad učenia optimálnej konfigurácie. ($\lambda=0.1, \mu=0, g=0.1$, epocha=1000, skrytá vrstva=19)

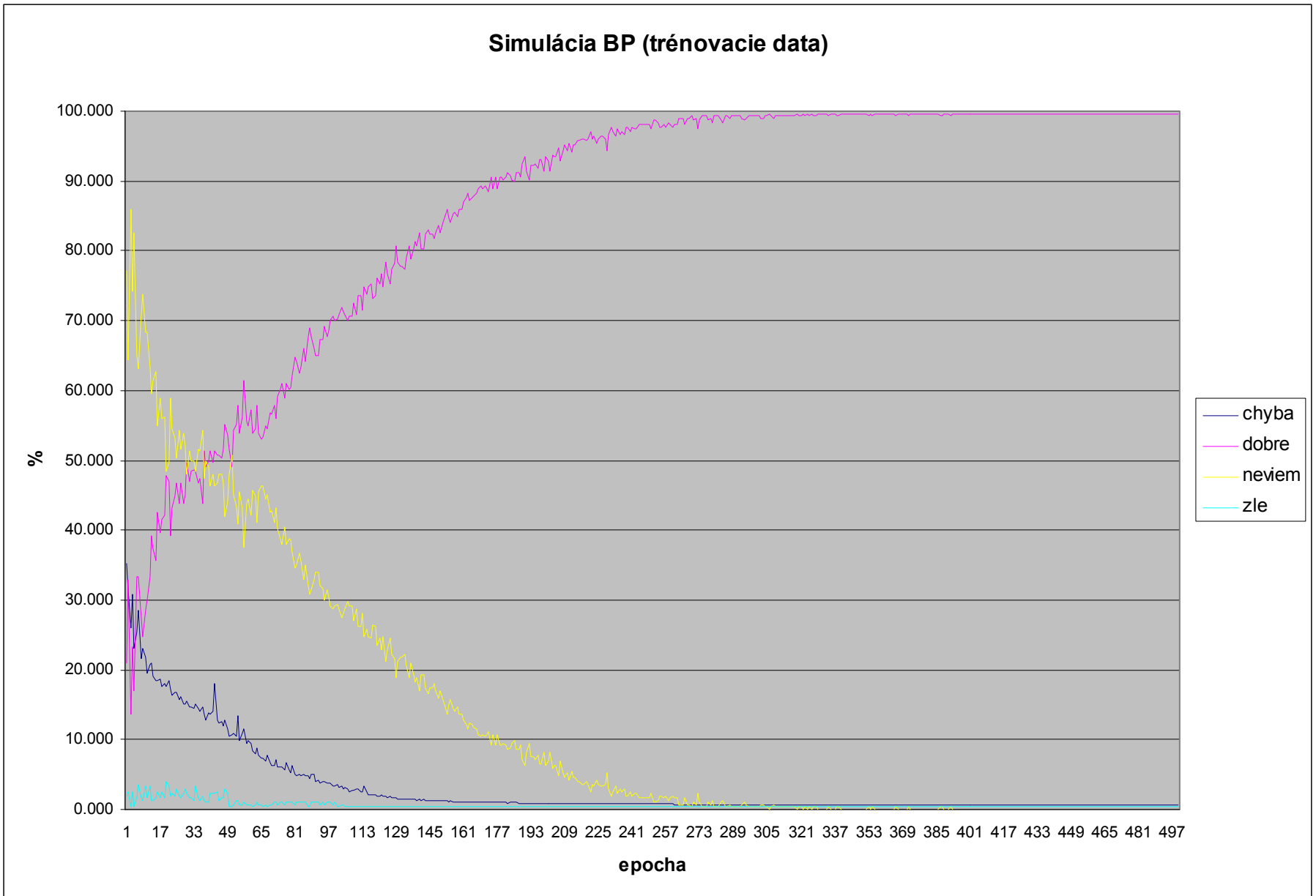
pokús	trénovacie dáta				testovacie dáta			
	Chyba	dobré	neviem	zle	chyba	dobré	neviem	zle
1	0.531135	99.67	0.00	0.33	3.646442	84.62	7.69	7.69
2	0.547632	99.67	0.00	0.33	4.147667	84.62	6.41	8.97
3	0.541120	99.67	0.00	0.33	4.642883	79.49	11.54	8.97
4	2.534176	98.37	0.00	1.63	3.213137	82.05	11.54	6.41
5	0.544197	99.67	0.00	0.33	4.714301	80.77	8.97	10.26

Tab5 - Prehľad učenia optimálnej konfigurácie. ($\lambda=0.1, \mu=0, g=0.2$, epocha=1000, skrytá vrstva=19)

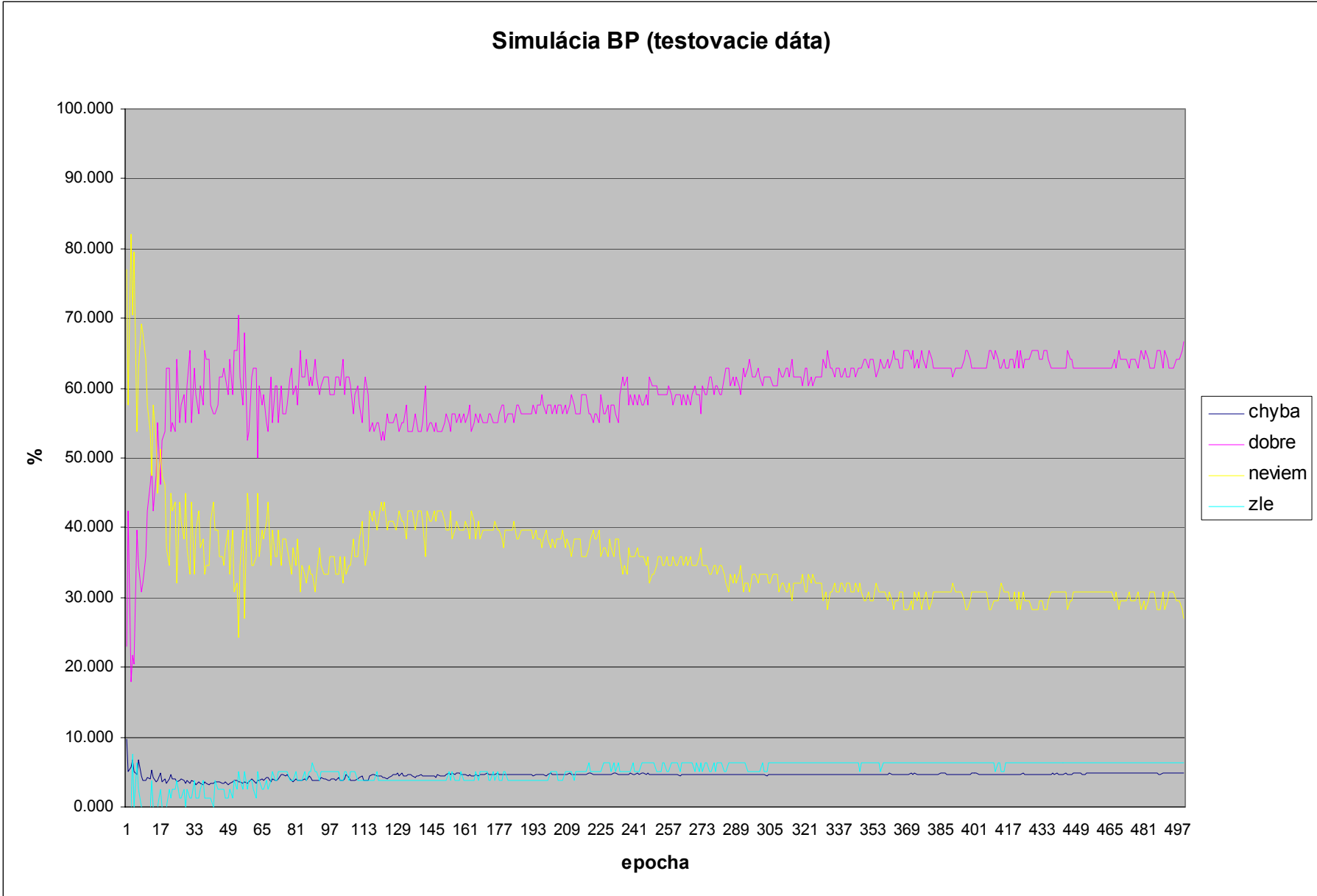
Porovnanie výstupu siete s trénovacou a testovacou množinou prikladám v **results.xls** (list simulate cmp).

Priebeh simulácie trénovanie siete s optimálnou konfiguráciou je znázornený v **(Graf1)** a jej testovania v **(Graf2)**.

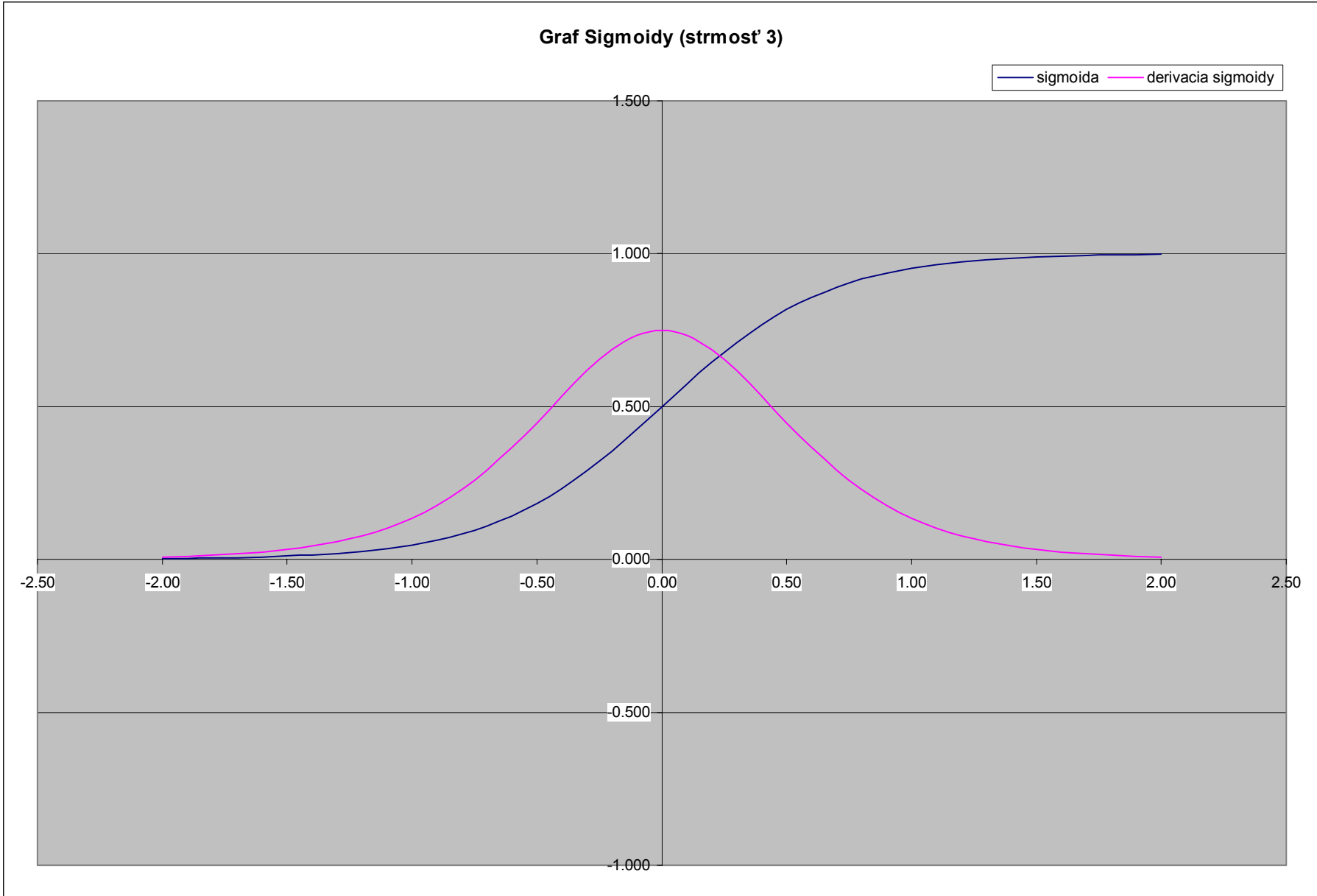
Program je v C++, vyvinutý v Dev-CPP.



Graf1 - simulácia učenia siete s optimálnou konfiguráciou (trénovacia množina)



Graf2 - simulácia učenia siete s optimálnou konfiguráciou (testovacia množina)



Graf3 - Sigmoida a jej derivácia