

Dáta

Zo súboru *omni2_24739.lst.txt*, ktorý obsahuje dáta zo všetkých rokov som pomocou jednoduchého skriptu (*preprocessingData.m*) vybral tie dáta, ktoré boli určené na tréningovanie, rok 2005, a tie, ktoré boli určené na testovanie, rok 2008. Tento program tiež vylúčil riadky s chýbajúcimi hodnotami (indikované ako 999.9, čo spôsobovalo problém pri normalizácii). Vznikli tak 2 súbory, a to *trainingYear2005.txt* a *testingYear2008.txt*.

Pomocou ďalšieho programu (*trainStorms.m*) som vyhľadal v tréningových rokoch búrkové dáta (146 = 36 + 2 + 108 hodín pre každú búrku), tieto hodnoty som znormalizoval do intervalu [-1; 1] a uložil znormalizované dáta do súboru *trainStormsNorm.txt*.

Model A

Prvým modelom, ktorý som sa pokúšal naučiť predpovedať geomagnetické búrky bol „feed-forward“ model A.

Celý beh (tréningovanie a testovanie) som spúšťal zo skriptu *main.m*. Tento skript spustil tréningovanie siete, znormalizoval dáta, spustil testovanie natréinovanej siete a vyhodnotil výsledky.

Tréningovanie

Na tréningovanie bol použitý klasický back-propagation algoritmus.

Tréningovanie prebiehalo na tréningových dátach v 80 cykloch. Počas tréningovania som experimentoval s počtom skrytých vrstiev (vyskúšal som 1, 2, 3, 4 a 5 skrytých vrstiev) a taktiež s počtom neurónov v nich. Tréningovanie tiež prebehlo s rôznymi počiatočnými váhami prepojení (intervaly [-2;2], [-1; 1], [-0.5; 0.5], [-0.25; 0.25], [0; 1], [0; 0.5], [0; 1/#vstupnýchNeurónov]) a rôznymi hodnotami eta-učiaceho pomeru (0.1 až 0.5).

Za finálnu topológiu siete som napokon zvolil: vstupná vrstva (41 neurónov), dve skryté vrstvy (33 a 9 neurónov) a výstupná vrstva (1 neurón). Vstupná aj skryté vrstvy obsahujú jeden „fiktívny“ neurón, ktorý vždy vracia hodnotu -1.

Za učiaci pomer som zvolil hodnotu 0,1 a váhy boli na začiatku stanovené na náhodné hodnoty z intervalu [-0.25; 0.25].

Tréningovanie vykonal program *trainNeuralNetworkA.m*.

Testovanie

Po natréningovaní spustil program *main.m* normalizáciu testovacích dát a napokon testovanie siete modelu A.

Testovanie vykonal program *testNeuralNetworkA.m*, ktorý dostal na vstup natréinovované váhy z tréningového cyklu a na výstup vrátil hodnoty D_{st} pre testovací rok.

Vyhodnotenie

Po skončení testovania som výsledky vyhodnotil nasledovne:

1. Vypočítal som strednú kvadratickú odchýlku, podľa vzorca $s = \sqrt{\frac{1}{T} \sum_{t=1}^T (d(t) - y(t))^2}$. Kde T je celkový počet hodín v danom roku, d je skutočná a y vygenerovaná hodnota D_{st} v hodine t .

2. Porovnal som reálny výskyt búrok v testovacích dátach s výskytom búrok v dátach predpovedaných sieťou.

Záver

Keďže sa jedná o feed-forward sieť, kde sieťou predpovedaná hodnota v bode i nie je závislá od hodnôt v bodoch $i-1$ až $i-t$, no v reálnej situácii (geomagnetické búrky) tomu tak je, predpokladám, že tento model nie je vhodný na túto úlohu. Druhou možnosťou je samozrejme mnou zlé nastavenie počtu skrytých vrstiev, počtu neurónov v nich, počiatkové nastavenie váh alebo učiaceho pomeru.

Model B

Druhým zvoleným modelom je model B, ktorého vstupná vrstva je oproti modelu A doplnená o 8 neurónov, ktoré prijímajú výstupné hodnoty siete spätným prepojením.

Celý beh (trénovanie a testovanie) som spúšťal zo skriptu *main2.m*. Tento skript spustil trénovanie siete, znormalizoval dáta, spustil testovanie natrénovanej siete a vyhodnotil výsledky.

Trénovanie

Keďže spätné prepojenie z výstupu siete je pevne nastavené na hodnotu 1, bolo možné použiť klasický back-propagation algoritmus.

Trénovanie prebiehalo v 100 cykloch, pričom trénovanie prebiehalo na jednotlivých búrkach v rámci jedného cyklu osobitne. Tento rozdiel oproti modelu A som implementoval z dôvodu, že model B pracuje s už vygenerovanými hodnotami v predchádzajúcom kroku a teda miešanie dát na rozhraní dvoch búrok môže spôsobovať nesprávne výsledky.

Pri trénovaní tejto siete som opäť experimentoval s počtom skrytých sietí a neurónov v nich, inicializačnými hodnotami váh, hodnotou η aj počtom cyklov.

Za finálnu topológiu siete som napokon zvolil: vstupná vrstva (49 neurónov), dve skryté vrstvy (33 a 15 neurónov) a výstupná vrstva (1 neurón). Vstupná aj skryté vrstvy obsahujú jeden „fiktívny“ neurón, ktorý vždy vracia hodnotu -1.

Za učiaci pomer som zvolil hodnotu 0,1 a váhy boli na začiatku stanovené na náhodné hodnoty z intervalu $[-0.5; 0.5]$.

Trénovanie vykonal program *trainNeuralNetworkB.m*.

Testovanie

Po natrénovaní spustil program *main2.m* normalizáciu testovacích dát a napokon testovanie siete modelu B.

Testovanie vykonal program *testNeuralNetworkB.m*, ktorý dostal na vstup natrénované váhy z tréningového cyklu a na výstup vrátil hodnoty D_{st} pre testovací rok.

Vyhodnotenie

Vyhodnotenie prebiehalo totožne ako vyhodnotenie modelu A, a to pomocou strednej kvadratickej odchýlky a porovnania výskytu búrok.

Záver

Keďže sa jedná o čiastočne rekurentnú sieť, predpokladal som „lepšie výsledky“ ako podal model A. Avšak vzhľadom na fakt, že reálne hodnoty D_{st} môžu v dôsledku silnej aktivity slnka „skákať“, je možné, že ani model B nedokázal tieto náhle zmeny reálne predpovedať. Druhou možnosťou je

samozrejme mnou zlé nastavenie počtu skrytých vrstiev, počtu neurónov v nich, počiatkové nastavenie váh alebo učiaceho pomeru.

Reálny výskyt búrok v testovacom roku 2008	
Začiatkový deň	Začiatková hodina
66	16
83	23
281	11

Výskyt búrok v tréningovom roku 2005	
Začiatkový deň	Začiatková hodina
10	14
47	11
64	1
126	11
138	16
148	9
162	5
189	22
234	20
242	0
251	2
296	16