

Samoorganizujúce sa neurónové siete s viacerými úrovňami laterálnych váh

Gabriela Andrejková, Jozef Oravec

Ústav informatiky, Prírodovedecká fakulta UPJŠ v Košiciach
Jesenná 5, Košice
Email: gabriela.andrejko@upjs.sk, jozi.oravec@gmail.com

Abstrakt

Grafy sú veľmi dôležitými modelmi dátových štruktúr, pretože poskytujú možnosť reprezentovať fakty vo vrcholoch a vzťahy medzi nimi pomocou orientovaných alebo neorientovaných hrán. V príspevku je spracovaný novo navrhnutý model samoorganizujúcej sa neurónovej siete (SOM NS) s viacerými vrstvami laterálnych váh, ktoré sa dokážu natrénovať a rozpoznávať grafové dátové štruktúry. Navrhnutý model bol implementovaný a vyhodnotený z hľadiska kvality získaných výstupných štruktúr. Konkrétne výsledky testov boli získané na štruktúrach usporiadania predmetov a ich prerekvizít v študijných programoch na univerzite. Predmety a ich prerekvizity sú usporiadané do acyklických grafových štruktúr. Experimenty na dátových štruktúrach tohto typu a vyhodnotenie výstupných grafových štruktúr z natrénovanej siete boli urobené z hľadiska kritéria usporiadania potomkov vrcholov výstupnej štruktúry. Vo vyhodnotení je pozorovateľné podstatné zlepšenie výsledkov už pri použití dvoch vrstiev laterálnych váh, použitie ďalších vrstiev prináša ďalšie zlepšenia.

1 Úvod

Objekty z reálneho sveta a vzťahy medzi nimi je možné reprezentovať pomocou grafov. Grafy sú veľmi dôležitými modelmi dátových štruktúr, pretože umožňujú reprezentovať fakty vo vrcholoch a vzťahy medzi nimi pomocou orientovaných hrán. Reprezentácia grafu pomocou matice (susednosti, incidencie, apod.) je možná a dobrá, avšak pre učenie z príkladov vzhľadom na rôzne veľkosti pre rôzne grafy je nevhodná. Preto je skôr používaná reprezentácia pomocou vymenovania hrán, pričom je možné použiť aj ohodnotenie hrán. Algoritmus učenia sa (trénovania) z príkladov musí toto zohľadňovať, a tak je jeden graf učení vo viacerých krokoch.

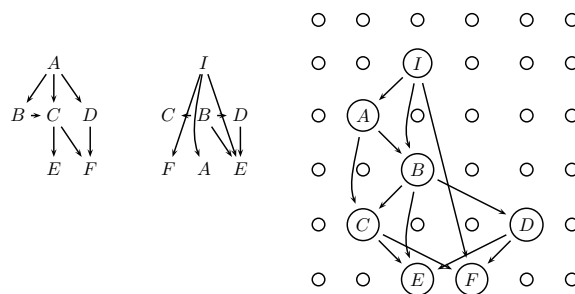
Samoorganizujúce sa mapy (SOM) pre štruktúrované dáta (SOM-SD) je model neurónovej siete vhodný pre riešenie problémov pracujúcich s grafovými štruktúrami a ich použitím sa zaoberali viacerí autori, napríklad Frasconi a spol. (1998), Hagenbuchner a spol. (2003), Hagenbuchner a Tsoi (2004), Scarselli a spol. (2009), Vančo a Farkaš (2010). Základný model SOM má jednoduchú štruktúru a jednoduchý algoritmus učenia, jeho popis je možné nájsť v monografiách Kohonen (1989), Haykin

(1999) a Hertz a spol. (1991). V príspevku Andrejková a Oravec (2012) boli popísané teoretické východiská pre aplikácie používajúce SOM model na prácu s acyklickými dátovými štruktúrami.

Článok je rozdelený do piatich sekcií. V druhej sekcii je spracovaný navrhnutý model samoorganizujúcej sa mapy pre štruktúrované dáta s viacerými vrstvami laterálnych váh, v tretej sekcii sú uvedené kritéria, podľa ktorých sú vyhodnotené z natrénovanej neurónovej siete získané výsledné dátové štruktúry, vo štvrtej sekcii je vyhodnotený experiment na reálnych dátach (študijné programy na univerzite ako acyklické dátové štruktúry) podľa navrhnutých kritérií. V závere sú zosumarizované a vyhodnotené dosiahnuté výsledky.

2 Viaceré úrovne laterálnych váh v SOM-SD

Ak pri trénovaní neurónovej siete je použitých viacero dátových štruktúr, môže nastať situácia, že sa dátové štruktúry nebudú dať správne rozpoznať, t.j. výsledné štruktúry získané rekurzívnou aktiváciou neurónov od víťazného neurónu nemusia zodpovedať očakávanej štruktúre.



Obr. 1: Príklad trénovania grafových štruktúr v SOM NS modeli s laterálnymi váhami. V ľavej časti sú dve grafové štruktúry a v pravej časti je príklad natrénovanej mriežky pre uvedené grafové štruktúry, pričom orientované hrany vyjadrujú laterálne váhy.

Na Obr. 1 je príklad dvoch dátových štruktúr spolu s ich možným natrénovaním v mriežke neurónovej siete ako je uvedené v pravej časti obrázku. Písmená vo vrcholoch štruktúr predstavujú diskretné ohodnotenia vrcholov a týmto ohodnoteniam po natrénovaní odpovedajú neuróny mriežky. Po natrénovaní môžu štruktúry v neuró-

novej sieti splynúť s časťami iných štruktúr a tým pádom nie je možné očakávať správne výsledné štruktúry od neurónovej siete.

Preto sme navrhli doplniť neurónovú sieť o laterálne váhy uložené vo viacerých úrovniach v rámci jednej mriežky. Tým sa docielí to, že po natrénovaní bude hľadaná najpodobnejšia štruktúra k vstupnej štruktúre pomocou úrovni laterálnych váh. Túto sieť nazveme sieť s laterálnymi váhami v úrovniach (SLVU).

Budeme používať nasledujúce označenie:

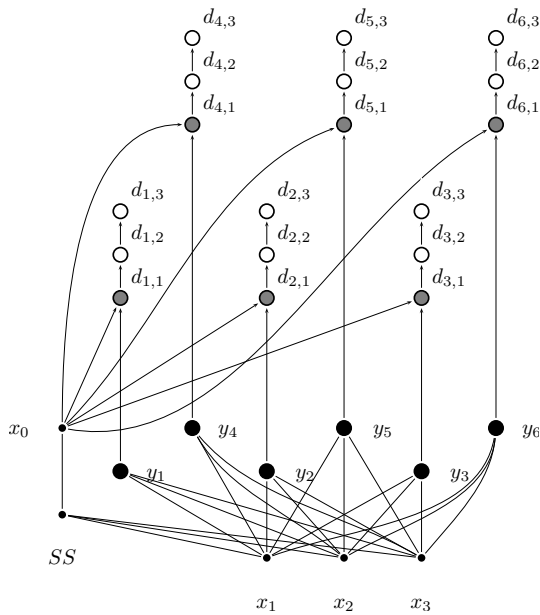
Graf G je možné vyjadriť ako dvojicu $G = (V, E)$, kde V je konečná množina vrcholov (fakty) $|V| = n$ a $E = \{(v, u); v, u \in V\}$ množina hrán (vzťahy) $|E| \leq n^2$. $ch_{\mathbf{D}}[v]$ predstavuje množinu všetkých priamych následovníkov vrcholu v v štruktúre \mathbf{D} . Definičný obor návěstí vrcholov grafu označíme \mathcal{U} a množinu všetkých dátových štruktúr v tomto obore $\mathcal{U}^\#$.

2.1 Organizačná dynamika SLVU

Naším cieľom je trénovať neurónovú sieť na acyklických grafových štruktúrach. Grafy majú ohodnotené vrcholy návěstiami (režazce alebo čísla). Trénovacia množina $\mathcal{T}^{\mathcal{L}}$ (1) je daná vstupnými štruktúrami \mathcal{U} z domény $\mathcal{U}^\#$. Bude použitý učiaci (trénovací) algoritmus bez dozoru (unsupervised learning).

$$\mathcal{T}^{\mathcal{L}} = \{\mathcal{U}; \mathcal{U} \in \mathcal{U}^\#\}, \quad (1)$$

V sieti sa vyžaduje nastavenie asociovanej vrstvy pre každú vstupnú dátovú štruktúru. Laterálne váhy v asociovanej vrstve budú adaptované pomocou vstupných dátových štruktúr.



Obr. 2: Neurónová sieť s 3 vrstvami laterálnych váh, 3 vstupnými neurónmi a mriežkou 3×2 , SS predstavuje rázcestník.

Na Obr. 2 je uvedený príklad neurónovej siete s mriežkou neurónov, nad mriežkou je niekoľko vrstiev laterálnych váh. Vľavo dole sa nachádza vstup x_0 s informáciou o asociovaných vrstvách. Vstupné neuróny sú x_1, x_2, x_3 . Ich aktivita zodpovedá návěstiam vrcholov v trénujúcej štruktúre. Časť s označením SS sa nazýva **rázcestník** siete a reprezentuje sieť SOM SD s počtom neurónov zodpovedajúcim počtu vrstiev laterálnych váh. Neuróny v prvej vrstve sú zakreslené sivou farbou, ďalšie vrstvy neurónov sú znázornené nevyplnenými krúžkami. Neuróny mriežky siete SOM sú znázornené čiernymi krúžkami.

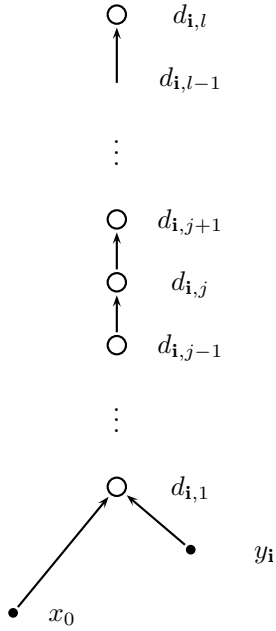
Priradenie niektorej vrstvy k dátovej štruktúre je vykonávané pomocou rázcestníka – SOM-SD siete uvedenej v Andrejková a Oravec (2012); mriežka rázcestníka je jednorozmerná a počet neurónov v nej sa rovná počtu vrstiev. Každý dátový štruktúre zodpovedá niektorý víťazný neurón a jeho index definuje vrstvu laterálnych váh.

Základ organizačnej dynamiky siete SLVU môže byť opísaný pomocou laterálnych váh a SOM SD takto:

1. Vstup $\mathbf{x}_I = (x_1, \dots, x_n)$ má n vstupných neurónov.
2. Ku vstupnej vrstve je pripojený rázcestník – jednorozmerná sieť SOM-SD s l neurónmi a váhami w_{ji}^R spájajúcimi j -tý vstupný neurón a i -tý neurón siete SOM-SD.
3. Nad vstupnou vrstvou sa nachádza spravidla 2-rozmerná mriežka s $N = N_1 \cdot N_2$ neurónmi, kde $N_j, j = 1, 2$ je počet neurónov v j -tom rozmere (je možné použitie viacrozmerých mriežok). Váha medzi i -tým neurónom siete a j -tým vstupným neurónom je w_{ji} . Neurón mriežky má index i a jeho stav je y_i . Mriežka patrí sieti SOM-SD.
4. V sieti sa nachádza l vrstiev laterálnych váh, počet vrstiev sa rovná počtu neurónov v rázcestníku, l je voliteľný parameter. Neuróny vo vrstvách sú prepojené laterálnymi váhami w_{ji}^s , kde $j, i = 1, \dots, N$ sú indexy neurónov v mriežke a s je index vrstvy.

Na Obr. 3 je znázornený vertikálny pohľad na usporiadanie neurónov. Nad i -tým neurónom v mriežke sa nachádzajú neuróny $d_{i,1}, \dots, d_{i,l}$, spojené váhami s hodnotou 1, označíme ich $w_{d_{i,s}, d_{i,s+1}}$. Laterálne váhy medzi neurónmi vo vrstve s sú w_{ji}^s vychádzajú z každého neurónu $d_{i,s}$ a vedú do neurónu $d_{j,s}$. Do neurónu $d_{i,1}$ vstupujú dve váhy, váha s hodnotou 1 z neurónu x_0 , ktorá určuje vrstvu (váha w_{i0}) a druhá váha w_i vychádza z neurónu i mriežky SOM siete, ktorá posúva aktivitu neurónu z mriežky do odpovedajúcej vrstvy laterálnych váh.

Medzi každým neurónom mriežky rázcestníka a neurónom odpovedajúcim vstupu x_0 je l váh s hodnotami $w_{s0} = s$, kde $s = 1, \dots, l$ je s -tý neurón rázcestníka. Rázcestník nasmeruje výpočet do s -tej vrstvy. Rázcestník má tiež kontextovú vrstvu s neurónmi c_1, \dots, c_o , kde $o = \max\{o_D; \mathbf{D} \in \mathcal{T}^{\mathcal{L}}\}$, o_D je maximálny počet



Obr. 3: Vertikálny pohľad na sieť SLVU s laterálnymi váhami vo vrstvách odpovedajúcich neurónu so súradnicami i v mriežke.

nasledovníkov vrcholov štruktúry \mathbf{D} . Neuróny kontextovej vrstvy sú úplne prepojené s jednorozmernou mriežkou a váhy medzi vrstvami sú w_{ji}^{RC} , kde $j = 1, \dots, o$, $i = 1, \dots, s$.

2.2 Adaptívna dynamika

Adaptívna dynamika spočíva v nastavení konfigurácie siete a vychádza z adaptívnej dynamiky siete SOM-SD. Pozostáva z dvoch krokov. V prvom kroku je trénovaný rázcestník. Rázcestník je trénovaný ako SOM-SD. Druhým krokom je adaptácia váh medzi mriežkou a vstupnou vrstvou a laterálnych váh v úrovniach. Počas adaptívnej dynamiky pri spracovaní jedného vrcholu v vstupnej štruktúry \mathbf{D} bude aktivovaný neurón y_{i^*} odpovedajúci víťazovi i^* na hodnotu 1, ktorého aktivita sa postupne posunie do s -tej úrovne a teda aktivita neurónu d_i^s bude 1. Cieľom je aktivovať neuróny d_k^s odpovedajúce potomkom vrcholu v na hodnoty $d_k^s(t) = a$, $a \in \langle \phi, 1 \rangle$, kde neurón k odpovedá víťaznému neurónu a ohodnoteniu vrchola $ch_{\mathbf{D},k}[v]$ počas iterácie t . ϕ je voliteľný parameter aktivácie neurónov. Úprava váh sa uskutoční naraz pre váhy medzi vstupnou vrstvou a mriežkou a laterálnymi váhami v príslušnej úrovni. Adaptácia váh j -tého neurónu je urobená podľa (2).

$$\Delta \mathbf{w}_j(t) = \eta(t) h_{i^*}(j, t) \|\mathbf{x} - \mathbf{w}_j(t-1)\|^2, \quad (2)$$

kde $\eta(t)$ je učiaci pomer, h_{i^*} je funkcia susedstva. Adaptácia laterálnych váh vo vrstve s je urobená nasledujúcim spôsobom:

$$hh = h^l(i_{\mathbf{D}_v}^*, i, j, x_j(t) - w_{ji}^s(t-1))$$

$$w_{ji}^s(t+1) = w_{ji}^s(t) + \eta(t) * hh * (x_j(t) - w_{ji}^s(t-1)) \quad (3)$$

kde $h^l(i_{\mathbf{D}_v}^*, j, t, p)$ je funkcia susedstva (4) a t je poradové číslo iterácie. Funkcia susedstva v úrovniach je vyjadrená v (4)

$$h^l(i_{\mathbf{D}_v}^*, i, j, p) = \begin{cases} \exp\left(-\frac{\|i - i_{\mathbf{D}_v}^*\|}{2\sigma^2(t)}\right), & 1 \leq i \leq N, 1 \leq j \leq n \\ \frac{1}{1+e^{-\alpha p}}, p \geq 0, & i = i^*, n+1 \leq j \leq n+N \\ \frac{1}{1+e^{-\beta p}}, p < 0, & i \neq i^*, n+1 \leq j \leq n+N \\ 0 & \text{inak,} \end{cases} \quad (4)$$

kde i^* je víťazný neurón, $\sigma(t) \in \mathfrak{R}$, $l = j - n$, $j \neq i$ je šírka susedstva, ktoré sa môže zmeniť počas učiaceho procesu, $p = x_j(t) - w_{ji}^s(t-1)$. Parametre α a β majú vplyv na excitáciu a na inhibíciu váh, $\alpha \geq 1$ a $\beta < 0$. Tento prístup je časovo náročný, pretože zmeny všetkých váh v sieti musia byť vždy znovu vypočítané.

Učiaci algoritmus - adaptácia SLVU:

Vstup: Trénovacia množina $T^{\mathcal{L}}$ definovaná v (1).

Výstup: Natrénované všetky váhy siete SLVU.

Metóda:

K0: Inicializovať náhodne laterálne váhy na úrovniach w_{ji}^s z intervalu $\langle 0, 1 \rangle$, inicializovať náhodné váhy w_{ji}^R , w_{ji}^{RC} , w_{ji} , váhy medzi rázcestníkom a neurónom x_0 nastaviť na $w_{s0} = s$, váhy $w_{d_{i,s}, d_{i,s+1}}$, w_i a w_{j0} nastaviť na 1.

K1: Aplikáciou algoritmu pre SOM-SD natrénovať rázcestník použitím trénujúcej množiny $T^{\mathcal{L}}$.

K2: Opakovať pre každú vstupnú dátovú štruktúru $\mathbf{D} \in T^{\mathcal{L}}$ a pre každý vrchol v inverzne usporiadanej množiny vrcholov $V_{inv}(\mathbf{D})$, pokiaľ sieť neskonverguje do stabilného stavu, nasledujúce kroky:

K2.1 Podľa rázcestníka určiť úroveň s pre štruktúru \mathbf{D} .

K2.2 Nájsť víťaza $i_{\mathbf{D}_v}^*$ pre ohodnotenie \mathbf{D}_v použitím (5).

K2.3 Zmeniť váhy w_{ji} podľa (2).

K2.4 Aktivovať neurón $d_{i,s}$ podľa (6).

K2.5 Ak má vrchol v potomkov, aktivovať neuróny vo vrstve s odpovedajúce potomkom z $ch_{\mathbf{D}}[v]$ tak, že ohodnotenie vrchola každého potomka bude dané na vstup a jeho aktivita z mriežky sa presunie na odpovedajúcu úroveň s podľa (6).

K2.6 Zmeniť váhu w_{ji}^s podľa (3).

K2.7 Deaktivovať všetky neuróny $d_{i,j}$ a y_i v mriežke, ktoré boli aktivované na hodnotu 0. Pokračovať spracovaním ďalšieho vrcholu.

2.3 Časová a priestorová zložitosť adaptácie

Pridanie vrstiev s laterálnymi váhami do siete sa odzrkadlí v priestorovej zložitosti. Teda, na vstupe je n neurónov spojených $n \cdot N$ váhami s mriežkou, v ktorej je N neurónov. V jednodimenzionálnej mriežke rázcestníka je l neurónov, ktoré sú spojené $n \cdot l$ váhami. Ďalej v rázcestníku je o kontextových neurónov s $o \cdot l$ váhami. Medzi rázcestníkom a neurónom x_0 je s váh. Medzi neurónom x_0 a prvou vrstvou laterálnych váh je N váh. Podobne medzi mriežkou a prvou vrstvou je N váh. V jednotlivých úrovniach je spolu $l \cdot N$ neurónov a $l \cdot N^2$ laterálnych váh. Medzi jednotlivými vrstvami je $(l - 1) \cdot N$ váh. Úhrnom dostávame odhad pre priestorovú zložitosť $S = O(n + l + o + s + (n + o + 2N + N^2) \cdot l + 2N)$.

Za predpokladu, že je natrénovaný rázcestník SS , odhad na tréning jednej vzorky má časovú zložitosť $T = O(nN + nN + s + 1 + so + N^2 + s(o + 1)) = O(so + 2nN + N^2)$, kde na nájdenie víťaza je potrebných nN operácií, zmena váh medzi mriežkou a vstupnými hodnotami si vyžiada nN operácií. Na presun aktivity víťazného neurónu do s -tej vrstvy je potrebných $s + 1$ operácií aj s presunom aktivity x_0 , na presun aktivít potomkov z s -tej vrstvy je potrebných $s \cdot |ch_D[v]| \leq s \cdot o$ operácií. Na počet zmien laterálnych váh treba N^2 operácií. Deaktivácia všetkých aktívnych neurónov bude vyžadovať najviac $s(o + 1)$ operácií.

2.4 Aktívna dynamika

Práca siete spočíva v rekurzívnej aktivácii neurónov použitím laterálnych váh vo vrstve určenej rázcestníkom. Najprv sú aktivované všetky vstupné neuróny hodnotami x_1, \dots, x_n . Nasledujú dva kroky: v prvom kroku je vypočítaný víťazný neurón v časti siete SS použitím klasického SOM algoritmu, pričom index víťazného neurónu určí vrstvu s ; v druhom kroku je určený víťaz mriežky i^* siete SOM podľa (5), kde \mathbf{x}_I sú stavy vstupných neurónov a \mathbf{w}_j^I sú váhy.

$$i_{D_v}^* = \arg \min_{j=1, \dots, N} (\|\mathbf{0D}_v - \mathbf{w}_j^I\|), \quad (5)$$

kde $\mathbf{D}_v = \mathbf{x}_I$ je ohodnotenie vrchola v , ktoré je na vstupe siete. Potom aktivita neurónu y_i v prvej vrstve je prenesená do vrstvy s nasledujúcim spôsobom: Stav neurónov $d_{i,j}$ budú vypočítané pre všetky vrstvy p takto:

$$d_{i,p} = \begin{cases} w_{i0}x_0 - 1 + w_i y_i & p = 1, \\ d_{i,p-1} - 1 & d_{i,p-1} > 1, 2 \leq p \leq s \\ d_{i,p-1} & d_{i,p-1} \leq 1, p > s, \end{cases} \quad (6)$$

kde $x_0 = \sum_{s=0}^l w_{s0} y_s^{SS}$.

Z hodnôt váh w_{s0} a z $y_1 = 1$ vyplýva, že neuróny vo vrstve s budú aktivované na hodnotu $d_{i,s} = 1$ (je to i -tý neurón v s -tej vrstve), aktivita neurónov zodpovedá víťaznému neurónu v mriežke prvej vrstvy.

¹Aktívny neurón má výstupnú hodnotu 1.

Neuróny vo vrstve s sú aktivované rekurzívne použitím laterálnych váh w_{ji}^s podľa algoritmu pre adaptáciu siete SLVU a tak sa dostane štruktúra do vrstvy trénovanej neurónovej siete. Použijeme voliteľný parameter ϕ - *dôležitosť prepojenia*, ktorým je možné regulovať aktiváciu neurónov. Ak $d_{i,j} \in \langle \phi, 1 \rangle$, neurón je aktivovaný, inak je neaktívny. Je potrebné poznamenať, že aj v tejto sieti môže dôjsť k *prekrytiu štruktúr*.

Poznámka: Výhodou vytvorenia niekoľkých úrovní mriežok sú lepšie sieťou vypočítané štruktúry, ale zásadná nevýhoda sa skrýva v rázcestníku, ktorý pozostáva zo siete SOM-SD. Aby bol získaný víťazný neurón v jednodimenzionálnej mriežke, je potrebné dať na vstup počas aktívnej dynamiky v prvej fáze výberu úrovne celú štruktúru, najlepšie tú, ktorá bola trénovaná, aby bola správne vybraná úroveň rázcestníkom. Nevýhodou tejto siete je aj to, že keď vieme len ohodnotenie vrchola, od ktorého chceme vedieť interpretáciu štruktúry sieťou, sieť SLVU nebude vedieť vybrať správnu úroveň, v ktorej sa má vykonať aktivácia neurónov.

3 Vyhodnotenie modelu SLVU

Označme T^O testovaciu množinu dátových štruktúr pre testovanie SLVU,

$$T^O = \{\mathbf{U}; \mathbf{U} \in \mathcal{U}^\#\} \quad (7)$$

Vrcholy v dátovej štruktúre sú označené jednoznačnými návěstiami, ich identifikácia je veľmi jednoduchá. Každý vrchol môže byť použitý vo viacerých štruktúrach. Vyhodnotenie je urobené detailným porovnaním štruktúry získanej výpočtom siete (sieťová interpretácia) a očakávanej (testovanej vstupnej) štruktúry.

3.1 Diferenciácia víťazov

Diferenciácia víťazov (DV) vyjadruje pomer medzi počtom víťazných neurónov a počtom rôznych vstupných dát Vančo a Farkaš (2010)

$$DV = \frac{|\{j; \exists t : j = i^*(t)\}|}{N}, \quad (8)$$

Ak $DV < 1$, tí istí víťazi sú vypočítaní pre viacej rôznych vstupov. Situácia je najlepšia v prípade, ak $DV = 1$. Predpokladáme, že parameter DV sa bude blížiť k 1.

3.2 Kvalita získaných dátových štruktúr v SLVU

Kvalita trénovanej siete SLVU bude vyhodnotená na základe zoznamu následovníkov vrcholov vstupnej štruktúry a sieťou vypočítanej štruktúry. Vstupná štruktúra je prezentovaná pomocou kódov návěstí na vstupe siete SLVU. Vyhodnotenie je urobené spracovaním aktivácií neurónov. Pred spracovaním prvého návestia na vstupe je vykonaná inicializácia neurónov v sieti a výstupnej

vrstvy ako neaktívne (hodnota 0). Po vykonaní výpočtu budú niektoré neuróny aktivované (hodnota 1). Aktivované neuróny zodpovedajú buď návštevu vrcholu na vstupe alebo návštevám jeho nasledovníkov. Aktivované neuróny predstavujú sieťou vypočítanú štruktúru a nemusia zodpovedať reálnej situácii. **Správna aktivácia neurónu** predstavuje neurón v stave 1 a jeho očakávaná hodnota je tiež 1. **Nesprávna aktivácia neurónu** predstavuje neurón v stave 1 a podľa vstupnej štruktúry by mal byť neaktívny, mal by mať hodnotu 0. Je možné použiť dva prístupy k vyhodnoteniu kvality vypočítanej štruktúry:

1. **dôveryhodnosť (SA)** vypočítanej štruktúry je pomer medzi počtom správne aktivovaných neurónov a počtom všetkých vrcholov v štruktúre, a
2. **nedôveryhodnosť (NA)** vypočítanej štruktúry je pomer medzi počtom nesprávne aktivovaných neurónov a počtom všetkých vrcholov v štruktúre.

Vyhodnotenie je počítané na testovacej množine pripravenej podľa (9).

$$0 = \{ \{ (\mathbf{U}_v, de_{\mathbf{U}}[v] \cup \{ \mathbf{U}_v \}) \} \}, \mathbf{U} \in \mathbf{T}^O, \quad (9)$$

kde $(\mathbf{U}_v, de_{\mathbf{U}}[v] \cup \{ \mathbf{U}_v \})$ sú dvojice návěstí vstupných vrcholov - vrchol a jeho nasledovníci a \mathbf{U} je dátová štruktúra z množiny štruktúr \mathbf{T}^O . Dôveryhodnosť SA a nedôveryhodnosť NA sú definované podľa (10):

$$SA = \frac{p^{SA}}{p^S}, \quad NA = \frac{p^{NA}}{p^S} \quad (10)$$

kde p^S je súčet počtov všetkých správne vymenovaných všetkých nasledovníkov vrcholov $v_{\mathbf{U}}$ z testovacej množiny 0 na vstupe, vrátane daného vrcholu, a p^{SA} je súčet všetkých správne aktivovaných neurónov (p^{NA} je súčet všetkých nesprávne aktivovaných neurónov) po sekvenčnom vstupe všetkých návěstí \mathbf{U}_v z testovacej množiny. $SA + NA = 1$, preto uvedieme výsledky pre dôveryhodnosť.

3.3 Kvalita topológie získaných dátových štruktúr

Výpočtom siete SLVU dostaneme vypočítanú dátovú štruktúru použitím rekurzívnej aktivácie, označme ju ${}_v\mathbf{U}^S$. Predpokladáme, že zodpovedá reálnej štruktúre ${}_v\mathbf{U}$. Použitím rekurzívnej aktivácie neuróny v sieti sú usporiadané podľa hodnôt laterálnych váh, t.j. následovníci nejakého vrcholu v ${}_v\mathbf{U}^S$ sú usporiadaní podľa hodnôt laterálnych váh. Budeme porovnávať kvalitu vypočítanej štruktúry ${}_v\mathbf{U}^S$ vzhľadom na reálnu štruktúru ${}_v\mathbf{U}$ porovnávaním následovníkov nejakého navzájom si odpovedajúceho vrcholu v oboch štruktúrach. Vzhľadom na možnosti akceptovania poradia potomkov niektorého vrcholu dostávame tri typy usporiadaní:

- **Usporiadanie určené sieťou (US):** V štruktúre vypočítanej neurónovou sieťou akceptujeme usporiadanie potomkov vrchola v poradí, v akom ho neurónová sieť vypočítala, t.j. v poradí podľa hodnoty laterálnych váh.

- **Relatívne usporiadanie (RU):** V štruktúre vypočítanej neurónovou sieťou neakceptujeme v usporiadaní tých potomkov už vybraného vrchola, ktorí nevyhovujú usporiadaniu voči skutočnosti, t.j. ak pre štruktúru ${}_v\mathbf{U}^S$ platí usporiadanie $(u, r) \prec_{{}_v\mathbf{U}^S} (u, t) \prec_{{}_v\mathbf{U}^S} (u, z)$, kde $r, t, z \in ch_{{}_v\mathbf{U}^S}[u]$ a v reálnej štruktúre ${}_v\mathbf{U}$ je usporiadanie hrán $(u, r) \prec_{{}_v\mathbf{U}} (u, z)$, kde $r, z \in ch_{{}_v\mathbf{U}}[u]$, tak vybrané vrcholy z potomkov pre porovnávanie budú len r, z .

- **Úplné usporiadanie (UU):** V štruktúre vypočítanej neurónovou sieťou akceptujeme len tých potomkov už vybraného vrcholu, ktorých usporiadanie je úplne totožné s očakávaným usporiadaním potomkov vrchola skutočnej štruktúry a navyše počet potomkov majú rovnaký. Konkrétne pre získanú štruktúru ${}_v\mathbf{U}^S$ pomocou neurónovej siete platí $(u, r) \prec_{{}_v\mathbf{U}^S} (u, t)$, kde $r, t \in ch_{{}_v\mathbf{U}^S}[v]$ a v reálnej štruktúre ${}_v\mathbf{U}$ je usporiadanie vrcholov $(u, r) \prec_{{}_v\mathbf{U}} (u, t)$, kde $r, t \in ch_{{}_v\mathbf{U}}[v]$ a je tiež splnené $|ch_{{}_v\mathbf{U}^S}[v]| = |ch_{{}_v\mathbf{U}}[v]|$.

Siete SLVU sú schopné vypočítať dátovú štruktúru spolu s jej topológiou. Po predložení ohodnotenia vrcholu na vstup siete mu bude odpovedať víťazný neurón. Rekurzívnu aktiváciu pomocou laterálnych váh sú aktivované neuróny, ktoré zodpovedajú ďalším vrcholom. Takto získame topológiu dátovej štruktúry z neurónovej siete.

Vyššie sme uviedli, že budeme pri vyhodnocovaní zohľadňovať tri spôsoby vyhodnotenia z pohľadu usporiadania potomkov získanej štruktúry. Preto pri jednotlivých parametroch porovnávania budeme uvádzať jedno z označení $X = \{US; RU; UU\}$ podľa typu hodnoteného usporiadania.

Dôveryhodnosť definovaná v (10) nadobudne tvar:

$$SA_R^X = \frac{F_{SA_R}^X}{N^R}, \quad (11)$$

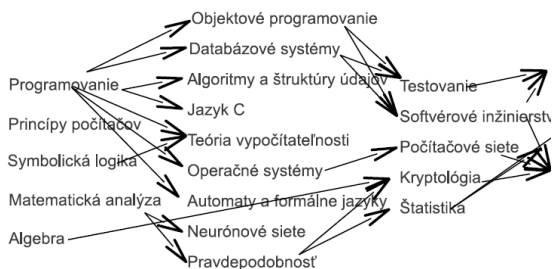
$$\text{kde } F_{SA_R}^X = \sum_{\mathbf{U} \in \mathbf{T}^O} 1 - \sum_{v \in V(\mathbf{U}) - \text{Listy}(\mathbf{U})} |SA_R^X(v)|,$$

a $SA_R^X(v)$ je počet správne rekurzívne aktivovaných neurónov zodpovedajúcim vrcholom – priamym potomkom $ch_{\mathbf{U}}[v]$, v zmysle vyššie uvedeného X , a N^R je počet všetkých vrcholov pri rekurzívnom spracovaní štruktúr, t.j. $N^R = \sum_{\mathbf{U} \in \mathbf{T}^O} (1 + \sum_v |ch_{\mathbf{U}}[v]|)$.

Kvalita SA_R^X vypočítanej dátovej štruktúry teda vyjadruje úspešnú aktiváciu neurónov odpovedajúceho vrcholu v a jeho priamych následovníkov, t. j. $de_{\mathbf{U}}[v] \cup \{v\}$.

4 Experiment: dátové štruktúry študijných programov

V mnohých študijných programoch (ŠP) na univerzitách sú používané tri typy predmetov: Povinné predmety, povinne voliteľné a výberové predmety. Z teoretického hľadiska všetky predmety tvoria čiastočne usporiadanú množinu podľa poradia pre ich absolvovanie. Čiastočne usporiadanie je určené prerekvizitami predmetov. Časť študijného programu Informatika na Prírodovedeckej fakulte UPJŠ je znázornená na Obr. 4.



Obr. 4: Časť študijného programu Informatika.

4.1 Príprava dátových štruktúr

Predmety študijných programov sú jednoznačne zakódované. Množina všetkých kódov je

$$P = \{\mathbf{p}^k; k = 1, \dots, K, \mathbf{p}^k \in \{0, 1\}^n\}, \quad (12)$$

kde $n = \log_2(2K + 10000) + 1$ a K je počet predmetov vo všetkých študijných programoch. Tieto kódy tvoria definičný obor návěstí dátových štruktúr.

Nech S je množina všetkých dátových štruktúr zodpovedajúcich študijným programom so spoločným štartujúcim vrcholom (supervrcholom):

$$S = \{\mathbf{D}; \mathbf{D} \text{ je dátová štruktúra jedného ŠP}\} \quad (13)$$

Označme $V(\mathbf{D})$ množinu všetkých vrcholov v dátovej štruktúre \mathbf{D} a $de_{\mathbf{D}}[v]$ množinu všetkých prerekvizitných vrcholov vrcholu $v \in V(\mathbf{D})$. Množina $V(\mathbf{D})$ je čiastočne usporiadaná podľa relácie **prerekvizita** - \prec a je vnorená do lineárneho usporiadania pred začiatkom tréningu. Usporiadanie je dôležité pre tréning a hlavne pre vyhodnotenie kvality SLVU.

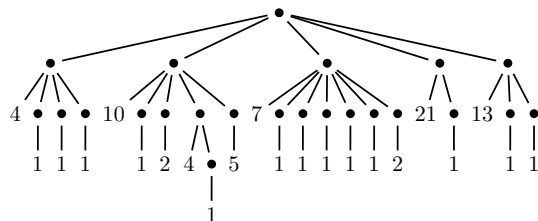
Experimenty boli urobené pre siete SLVU s 1 až 10 vrstvami laterálnych váh a pre štvorcové rozmery mriežky: $15 + 3k$, $k = 0, \dots, 13$. Siete boli tréňované vykonaním 3000 iterácií a ďalšie parametre boli nastavené na $\alpha = 2$, $\beta = -5.0173$, dôležitosť prepojenia $\phi = 0.95$.

Tréningová množina bola pripravená z 19 prvkov množiny S . Testovacia množina T^O pozostávala z podštruktúr prvkov množiny S .

$$T^O = \{v\mathbf{D}; v \in V(\mathbf{D}), \mathbf{D} \in T^{\mathcal{L}}\} \quad (14)$$

kde $v\mathbf{D}$ je graf nasledovníkov vrcholu v v štruktúre \mathbf{D} . V testovacej množine sa nachádzalo 5366 vrcholov vo všetkých úrovniach testovaných podštruktúr.

Príklad stromovej štruktúry z trénujúcej množiny štruktúr je uvedený na Obr. 5. Strom reprezentuje rekurzívny prechod acyklickej štruktúry odpovedajúcej študijnému programu bakalárskeho štúdia Informatika - Ib.



Obr. 5: Príklad trénujúcej štruktúry hĺbky 4: Štruktúra odpovedajúca študijnému programu Informatika - Ib.

4.2 Diferenciácia víťazov, výsledky

Úroveň	15	24	30	42	54
1	0,259	0,512	0,604	0,749	0,828
2	0,289	0,495	0,720	0,936	0,995
3	0,292	0,507	0,737	0,926	0,995
4	0,291	0,500	0,722	0,933	0,996
5	0,287	0,511	0,737	0,913	0,993
6	0,288	0,484	0,721	0,924	0,992
7	0,291	0,518	0,721	0,939	0,997
8	0,291	0,509	0,754	0,937	0,997
9	0,286	0,495	0,741	0,943	0,993
10	0,287	0,489	0,750	0,914	0,992

Tab. 1: Vyhodnotenie VD v sieti SLVU s počtom úrovní 1 až 10 a veľkosťou mriežky 15, 24, 30, 42, 54.

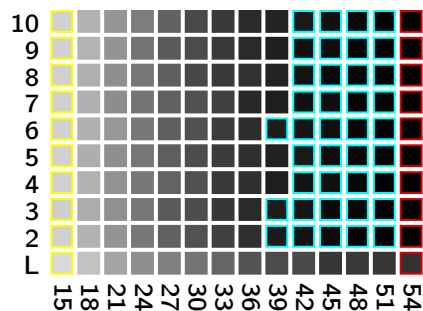
Najmenšia zistená hodnota je 0,259 pre rozmer mriežky 15 a maximálna je 0,997 pre rozmer mriežky 54. Z toho vyplýva, že ani v jednom prípade sa nedosiahol stav, v ktorom jednému neurónu mriežky odpovedá najviac jeden vrchol štruktúry.

Na Obr. 6 sú zobrazené graficky jednotlivé diferenciácie víťazov. Na grafe je vidieť, že hodnoty DV závisia len od rozmeru mriežky čo potvrdzujú aj údaje v Tab. 1.

V prípade sietí SLVU nie je použitý žiadny kontext. Na vstup sú kladené len ohodnotenia vrcholov a v prípade sietí s rozmerom mriežky od 42 hodnota diferenciácie je väčšia než 0.9, a teda môžeme pre tieto siete očakávať primerané hodnoty dôveryhodností.

4.3 SLVU - dôveryhodnosť, výsledky

Vyhodnotenie dôveryhodnosti siete SLVU v zmysle X je uvedené v Tab. 2. Minimálna hodnota dôveryhodnosti bola nameraná 0.2869 a maximálna hodnota 0.8942.



Obr. 6: Farebne odlišené hodnoty diferenciácie víťazov pre siete SLVU s mriežkami 15 až 54. Tmavšie hodnoty znamenajú lepšie dosiahnuté výsledky. Oramované štvorcečky predstavujú najlepšie/najhoršie hodnoty pre danú kombináciu počtu úrovní a veľkosti mriežky.

Vidíme, že ani v jednom prípade sieť nevedela vymenovať všetkých potomkov vrcholov štruktúr správne. Ak sa pozrieme na jednotlivé hodnoty, tak vždy najhoršie hodnoty boli pre siete s najmenšou mriežkou 15 a najlepšie hodnoty sa pohybovali na mriežkach 48, 51 a 54. Podľa očakávania najhoršie výsledky dosiahla sieť s jednou úrovňou laterálnych váh. Pridávaním úrovní sa výsledky vylepšovali. Najlepšie hodnoty dosiahla sieť so 6 úrovňami a rozmerom mriežky 54.

Trénovanie takýchto sietí je časovo aj pamäťovo náročné, a preto pre tak malý počet štruktúr v trénovacej množine T^L nie je potrebné používať siete s tak veľkým rozmerom a počtom úrovní.

Toto pozorovanie je možné vidieť aj v Tab. 2 aj v grafe na Obr. 7, na ktorom je viditeľné, že hodnoty dôveryhodnosti podľa očakávaného usporiadania X majú zásadné rozdiely. V prípade usporiadania daného sieťou a očakávaného relatívneho usporiadania sú hodnoty dôveryhodnosti veľmi podobné, čo znamená, že sieť, ak už vie vymenovať potomkov, tak dostatočne dobre zachová usporiadanie potomkov vrcholu voči skutočnej štruktúre. Horšie to je v prípade očakávaného úplného usporiadania. Tú sú hodnoty výrazne nižšie, čo je odôvodniteľné tým, že sieť dokáže podstatne ťažšie interpretovať štruktúru v pôvodnom usporiadaní.

5 Záver

V príspevku sme uviedli nový model SOM neurónových sietí s viacerými úrovňami laterálnych váh – SLVU model pre prácu s grafovými dátovými štruktúrami. Experiment s týmto modelom pri acyklických dátových štruktúrach predstavujúcich študijné programy na univerzite potvrdil použiteľnosť tohto modelu. Vyhodnotenie bolo urobené jednak z hľadiska usporiadania potomkov jednotlivých vrcholov, ale bol tiež sformulovaný problém *zápis predmetu s prerekvizitami*, v ktorom boli vyhodnotené odpovede systému pre jednotlivé predmety. Aj keď systém odpovedal väčšinou správne, jednalo sa len o formálne

V		15	30	45	51	54
1	ON	0,295	0,469	0,619	0,610	0,648
	RO	0,295	0,467	0,617	0,608	0,639
	FO	0,290	0,409	0,537	0,536	0,585
2	ON	0,293	0,460	0,718	0,829	0,869
	RO	0,292	0,458	0,716	0,821	0,864
	FO	0,287	0,408	0,617	0,659	0,678
3	ON	0,306	0,523	0,777	0,847	0,836
	RO	0,304	0,521	0,774	0,842	0,834
	FO	0,291	0,426	0,663	0,657	0,641
4	ON	0,325	0,581	0,805	0,854	0,884
	RO	0,324	0,578	0,802	0,849	0,881
	FO	0,296	0,445	0,665	0,708	0,789
5	ON	0,308	0,519	0,745	0,813	0,854
	RO	0,308	0,518	0,743	0,808	0,850
	FO	0,287	0,441	0,683	0,689	0,668
6	ON	0,359	0,677	0,866	0,860	0,894
	RO	0,358	0,676	0,862	0,857	0,892
	FO	0,314	0,543	0,654	0,763	0,757
7	ON	0,346	0,662	0,804	0,877	0,881
	RO	0,344	0,660	0,802	0,874	0,881
	FO	0,304	0,541	0,626	0,653	0,682
8	ON	0,362	0,685	0,833	0,884	0,871
	RO	0,360	0,681	0,827	0,878	0,868
	FO	0,316	0,535	0,715	0,647	0,706
9	ON	0,372	0,741	0,859	0,885	0,869
	RO	0,366	0,738	0,854	0,883	0,865
	FO	0,316	0,591	0,725	0,733	0,650
10	ON	0,355	0,610	0,811	0,858	0,844
	RO	0,352	0,608	0,807	0,851	0,841
	FO	0,306	0,530	0,631	0,680	0,675

Tab. 2: US, RU, UU - dôveryhodnosť. Stĺpce odpovedajú rozmeru mriežky, trojice riadkov odpovedajú hodnotám usporiadaní, kde uvedené číslo V vyjadruje počet úrovní. Vyhodnotenie je uvedené pre veľkosť štvorcovej mriežky 15, 30, 45, 51, 54.

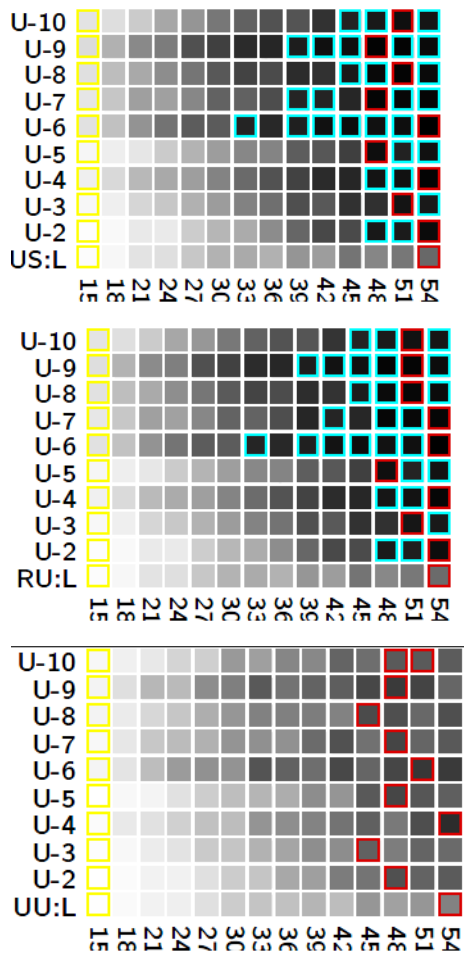
prerekvizity, čo pre študentov nie je postačujúce.

PodĎakovanie

Tento príspevok vznikol za podpory grantovej agentúry VEGA v rámci grantovej úlohy 1/0142/15.

Literatúra

- Andrejková, G. a Oravec, J. (2012). Application of modified som neural networks on acyclic data structures. *Acta Electrotechnica et Informatica*, 12(2):3–8.
- Frasconi, P. M., Gori, M. a Sperduti, A. (1998). A general framework of adaptive processing of data structures. *IEEE Transactions on Neural Networks*, 9(5):768–786.



Obr. 7: Farebne odlišené namerané hodnoty dôveryhodnosti pre siete SLVU s počtom úrovní 1 až 10 a s mriežkami 15 až 54 cez všetky úrovne štruktúr. Tmavšie sfarbené štvorčeky znamenajú lepšie namerané hodnoty. Oramované štvorčeky predstavujú najlepšie/najhoršie hodnoty pre danú kombináciu počtu úrovní a veľkosti mriežky.

Hagenbuchner, M., Sperduti, A. a Tsoi, A. C. (2003). A self-organizing map for adaptive processing of structured data. *IEEE Transactions on Neural Networks*, 14(3):491–505.

Hagenbuchner, M. a Tsoi, A. C. (2004). A supervised self-organizing map for structures. *Proceedings IEEE International Joint Conference on Neural Networks*, 14:1923–1928.

Hammer, B., Micheli, A., Sperduti, A. a Strickert, M. (2004). A general framework for unsupervised processing of structured data. *Neurocomputing*, (57):3–35.

Haykin, S. (1999). *Neural networks: a comprehensive foundations*. Prentice-Hall, New Jersey, 2. vyd.

Hertz, J., Krogh, A. a Palmer, R. (1991). *Introduction to the Theory of Neural Computation*. Addison-Wesley.

Kohonen, T. (1989). *Self-organization and associative memory*. Springer-Verlag, New York, NY, USA, 3. vyd.

Scarselli, F., Gori, M., Tsoi, A. C., Hagenbuchner, M. a Monfardini, G. (2009). The graph neural network model. *IEEE Transactions on Neural Networks*, 20(1):61–80.

Vančo, P. a Farkaš, I. (2010). Experimental comparison of recursive self-organizing maps for processing tree-structured data. *Neurocomputing*, (73):1362–1375.