

22. slovenské geodetické dni

Žilina 6. - 7. novembra 2014

APLIKÁCIA UMELÝCH NEURÓNOVÝCH SIETÍ NA TVORBU PREDIKČNÝCH MODELOV V GIS

APPLICATION OF ARTIFICIAL NEURAL NETWORKS FOR CREATION OF PREDICTIVE MODELS IN GIS

Filip Jurčiak¹

Abstract

This contribution deals with issue application of artificial neural networks in geographic information system (GIS). The artificial neural network is represented as universal tool for creating a predictive models. This technique of making predictive models is not much extended in Slovakia, but in abroad it is relatively well known. In this contribution is artificial neural network applied to predict the risk of landslides in Trenčín and Žilina region. In pursuance of four different parameters (slope of relief, orientation of relief, land cover and subsoil geological structure) was created artificial neural network able to predict the risk of landslides. The result is predictive model risk of landslides.

Úvod

Problematika zosuvov je na území Slovenskej republiky stále aktuálnou témou. Dôvodom sú časté zosuvy v oblastiach spojených s výstavbou rodinných domov, alebo rôznej ďalšej výstavby. Dôležitým procesom sa tak stáva spoľahlivejšia predikcia zosuvov, aby sa tak predišlo nebezpečenstvu, vyplývajúcemu zo samotnej povahy zosuvov. Na predikciu zosuvov sú často používané špecifické predikčné modely, založené na vopred definovaných matematických vzťahoch. V súčasnosti sa vo svete začínajú uplatňovať modely tvorené pomocou umelých neurónových sietí. Takto vytvorené modely predstavujú veľmi kvalitnú pomôcku pri lokalizácii potencionálne nebezpečných zosuvných zón.

V tomto príspevku bude vysvetlený proces využitia umelých neurónových sietí na tvorbu predikčných modelov. Ich základom je napodobňovanie procesov, ktoré sa odohrávajú v ľudskom mozgu. Počiatky neurónových sietí sú spojené so snahou modelovať nervovú bunku začiatkom 50. rokov. S nástupom počítačov sa čoraz viac vyvíjali aj neurónové siete, ale ešte stále to boli len veľmi jednoduché umelé neurónové siete. V 80. rokoch minulého storočia sa

¹ Bc. Filip Jurčiak, Klačno 2000/25, 034 01 Ružomberok, tel.: +421915307221, e-mail: jurciakfilip@gmail.com

začali vyrábať čoraz viac výkonné počítače, a tým umožnili tvorbu komplikovanejších umelých neurónových sietí. Cieľom umelých neurónových sietí je riešenie problémov tak, ako by to riešil odborník v danom odbore. V poslednom období sú umelé neurónové siete dokonca aplikované aj v prostredí sociálnych sietí na internete. Spoločnosť Facebook plánuje využívať neurónovú sieť pri rozpoznávaní tvári používateľov na fotografiách. Modely predikujúce zosuvy vytvorené v zahraničí dosahujú až 90% presnosť pri validácii výsledkov. Problematike takto vytvorených modelov nie je u nás venovaná dostatočná pozornosť, tak ako sa to deje v zahraničí.

1 Neurónová sieť

Základným elementom umelej neurónovej siete je neurón, ktorý pracuje rovnako ako neurón v ľudskom mozgu. Neurón teda vysiela a spracováva signály. Neuróny sú medzi sebou pospájané väzbami.

Na začiatok stručne ozrejmíme čo je to zosuv. Zosuv svahu alebo zosun svahu je relatívne rýchla svahová deformácia spôsobená gravitáciou, pri ktorej dochádza k pohybu horninového pokryvu po šmykových plochách. Pod termínom zosuv sa tiež rozumie geomorfologický tvar vytvorený týmto procesom [1].

Základný postulát neurovedy hovorí, že základným stavebným kameňom ľudského mozgu je neurón. Neurón má orgány (dendrity), ktorými prijíma signály z okolia od ostatných neurónov. Ďalej neurón spracováva prijaté signály a má orgány, ktorými posiela spracované vstupné signály iným neurónom zo svojho okolia.

1.1 Proces učenia

Medzi základné pojmy v teórii neurónových sietí patrí pojem učenie. Vo všeobecnosti je učenie proces, pri ktorom dochádza ku získavaniu skúseností, rozširovaniu poznatkov a k trvalej zmene osobnosti. Pojem učenia v teórii umelých neurónových sietí je definovaný nasledovným spôsobom. Učenie je proces, kedy sa znalosti ukladajú do synaptických váh umelej neurónovej siete. Ak si označíme maticu W ako maticu všetkých synaptických váh umelej neurónovej siete, tak pod učením budeme chápať stav kedy platí

$$\frac{\partial W}{\partial t} \neq 0, \quad (1.1)$$

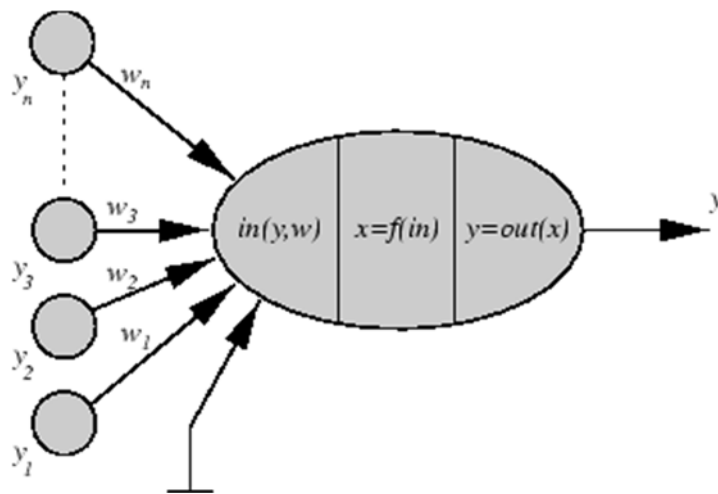
kde t je čas učenia. Teda synaptické váhy sa počas učenia menia [2].

Tento proces by sme mohli pomenovať aj adaptáciou umelej neurónovej siete. Jedná sa teda o zbieranie poznatkov a ich uchovanie. Proces učenia v zásadnej miere odlišuje umelé neurónové siete od ostatných výpočtových procesov. Pri nich je vždy potrebné navrhnuť nový algoritmus, pomocou ktorého by daný výpočet mohol prebehnúť. Neurónové siete majú navrhnutý svoj vlastný všeobecný algoritmus učenia a jeho vhodnosť determinuje kvalitu a rýchlosť učenia sa. Ako som už spomenul vyššie, neurónové siete sú založené na procese učenia. Pre tento proces je teda potrebné mať dáta, na základe ktorých sa môže neurónová sieť učiť simulovať nami skúmaný proces. Tieto dáta budeme nazývať reprezentatívnou vzorkou. Takúto vzorku je vhodné náhodne rozdeliť do dvoch základných typov vzoriek:

- a) Trénovacia vzorka predstavuje množinu hodnôt. Jej význam je kritický pre celý proces učenia. Dôvodom jej kritickosti je, že pri procese učenia sa používajú dáta práve z tejto vzorky. Tieto dáta sú extrahované do synaptických váh umelej neurónovej siete.
- b) Testovacia vzorka je taktiež množina usporiadaných hodnôt, ktorá sa používa v tzv. fáze života. Služí nám na otestovanie získaných znalostí počas učenia.

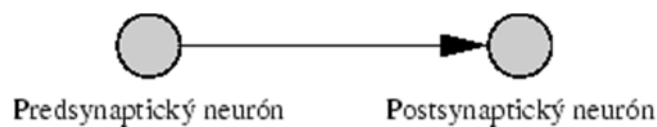
1.2 Neurón

Z predchádzajúceho vyplýva, že základom umelej neurónovej siete je samotný neurón. Stavbu neurónu (Obr. 1) môžeme popísať vstupom neurónu (dentrit), prahom neurónu θ_i , aktivačnou funkciou neurónu f , výstupnou funkciou neurónu o_i a synaptickými váhami w_i .



Obr. 1 Štruktúra neurónu [2]

Neuróny možno podľa toku signálu rozdeliť (Obr. 2) na predsynaptické (zdrojové alebo pred synapsiou) a postsynaptické (cieľové alebo po synapsii).



Obr. 2 Označenie neurónov [2]

Vstup do neurónu je uvažovaný ako funkcia jednotlivých vstupov z predsynaptických neurónov. Najčastejšie sa jedná o súčet takýchto vstupov s určitými váhami. Platí teda môže

$$in_i = \sum_{j=1}^N w_{i,j} ou_j + \theta_i, \quad (1.2)$$

kde N predstavuje počet predsynaptických neurónov, $w_{i,j}$ sú synaptické váhy, ou_j sú výstupy z neurónov a θ_i je prah i -teho neurónu.

Prah je vlastne vstup do neurónu z vonkajšieho sveta, teda nie z iných neurónov [2]. Môžeme teda povedať, že neuróny, ktoré nemajú žiadne vstupy z ostatných neurónov, majú vstup, ktorý predstavuje práve prah θ_i . Takéto neuróny sú nazývané sigma neuróny.

Aktivačná funkcia predstavuje stav neurónu v čase t respektíve v čase $t+1$. Možno ju zapísať nasledovne

$$x_i = f(in_i), \quad (1.3)$$

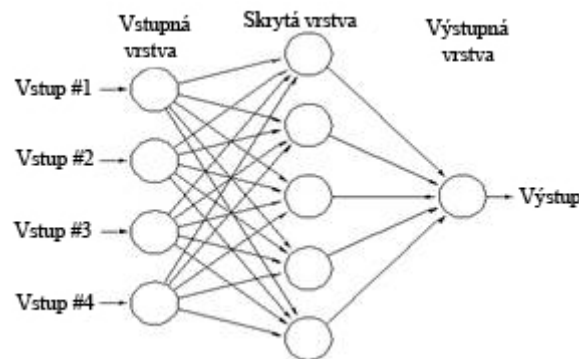
kde x_i je premennou a $f(in_i)$ je aktivačnou funkciou vstupu do neurónu. Tvar aktivačných funkcií môže byť rôzny.

Výstupná funkcia je často identickou aktivačnej funkcii a platí pre ňu

$$ou_i = O(x_i) = x_i = f(in_i). \quad (1.4)$$

1.3 Topológia umelej neurónovej siete

Neurónová sieť je vo všeobecnosti popísateľná orientovaným grafom (Obr. 3) pomocou vrcholov (neuróny) a orientovaných hrán (prepojenia). Popísať vlastnosti takýchto sietí je veľmi náročné. Preto sú analyzované siete s pravidelnou štruktúrou napríklad viacvrstvová štruktúra. Skladá sa zo vstupnej vrstvy, v ktorej neuróny dostávajú vstup len z vonkajšieho sveta a výstup obvykle pokračuje k ďalším neurónom. Ďalej sa skladá zo skrytej vrstvy, v ktorej do neurónov prichádzajú signály z ostatných neurónov, alebo externého sveta cez prahové spojenia a signály vysielajú ďalej. Výstupná vrstva je podobná skrytej vrstve s tým rozdielom, že výstup z tejto vrstvy ústi do externého sveta [3].



Obr. 3 Dopredná neurónová sieť [2]

Podľa vrstiev rozoznávame aj samotné neuróny. Poznáme vstupné, skryté a výstupné. Samotnú neurónovú sieť rozdeľujeme podľa topológie na dve základné skupiny:

- dopredné umelé neurónové siete (z angl. *feed-forward*),
- rekurentné umelé neurónové siete (z angl. *recurrent*).

1.4 Dopredné umelé neurónové siete

V takýchto umelých neurónových sieťach nie je možné jednoduchou linearizáciou systému rovníc popísať aktivity skrytých a výstupných neurónov. Dôležitým krokom je tu adaptácia

umelej neurónovej siete, ktorá minimalizuje účelovú funkciu. Medzi najefektívnejšie metódy minimalizácie patria tzv. gradientové metódy, založené na poznaní gradientu účelovej funkcie. Konkrétne sa jedná o parciálne derivácie vzhľadom k prahovým a váhovým faktorom

$$\frac{\partial E}{\partial w_{ij}} = \frac{\partial E}{\partial x_i} \frac{\partial x_i}{\partial w_{ij}} = \frac{\partial E}{\partial w_{ij}} t'(\xi_i) x_j, \quad (1.5)$$

$$\frac{\partial E}{\partial \vartheta_{ij}} = \frac{\partial E}{\partial x_i} \frac{\partial x_i}{\partial \vartheta_{ij}} = \frac{\partial E}{\partial x_{ij}} t'(\xi_i), \quad (1.6)$$

kde E je účelová funkcia a $t'(\xi_i)$ je prechodová funkcia.

Účelová funkcia vyjadruje rozdiel medzi výstupných hodnôt a požadovaných hodnôt.

Metóda najprudšieho spádu (*angl. steepest descent*) je v rámci optimalizačných metód jedným z najjednoduchších a najviac používaným spôsobov, ako realizovať túto minimalizáciu, v ktorej váhové a prahové koeficienty sú rekurentne obnovované v iteráciách pomocou vzťahov

$$w_{ij}^{(k+1)} = w_{ij}^{(k)} - \lambda \frac{\partial E}{\partial w_{ij}} + \mu \Delta w_{ij}^{(k)}, \quad (1.7)$$

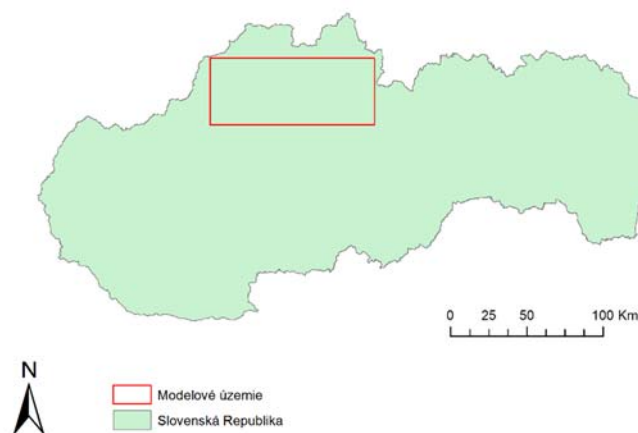
$$\vartheta_{ij}^{(k+1)} = \vartheta_{ij}^{(k)} - \lambda \frac{\partial E}{\partial \vartheta_{ij}} + \mu \Delta \vartheta_{ij}^{(k)}, \quad (1.8)$$

kde parameter $\lambda > 0$ je rýchlosť učenia a musí byť dostatočne malý (obvykle $\lambda = 0,01 - 0,1$), aby bola zabezpečená konvergentnosť optimalizačného algoritmu a súčasne dostatočne veľký pre zabezpečenie dostatočnej vysokej rýchlosti konvergentnosti.

Umelá neurónová sieť spomínaná v tomto príspevku bola vytvorená pomocou metódy (algoritmu) škálovaných konjugovaných gradientov (*angl. Scaled conjugate gradient (SCG)*) a metódu gradientovou (gradientovú (*angl. Gradient descent with momentum (GDM)*)).

2 Tvorba predikčného modelu

Predikčný model bol vytvorený na území Trenčianskeho a Žilinského kraja (*Obr. 4*). Táto oblasť je vhodná, pretože patrí medzi oblasti v minulosti z častým výskytom svahových deformácií a z vysokou pravdepodobnosťou výskytu ďalších svahových deformácií. Jedným z najznámejších zosuvov v tejto oblasti je Veľkomarský zosuv v tesnej blízkosti vodnej nádrže Liptovská Mara, s dĺžkou približne 1000 metrov a šírkou približne 600 m [4].



Obr. 4 Modelové územie

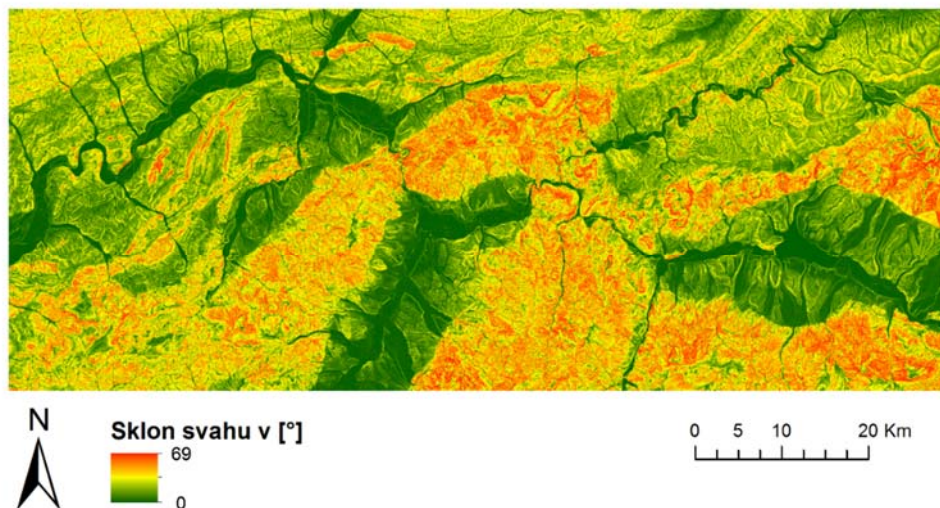
2.1 Parametre neurónovej siete

Vo všeobecnosti sú zosuvy náchylné na viacero parametrov. Napríklad v práci [5] boli na vytvorenie predikčného modelu použité parametre ako geologické pomery, nadmorská výška, sklon svahov, orientácia svahov, dĺžky svahov a prispievajúce plochy, krivosť reliéfu, krajinná pokrývka a samotné svahové deformácie. V uvedenej umelej neurónovej sieti sú využité údaje z nasledujúcich v podkladoch:

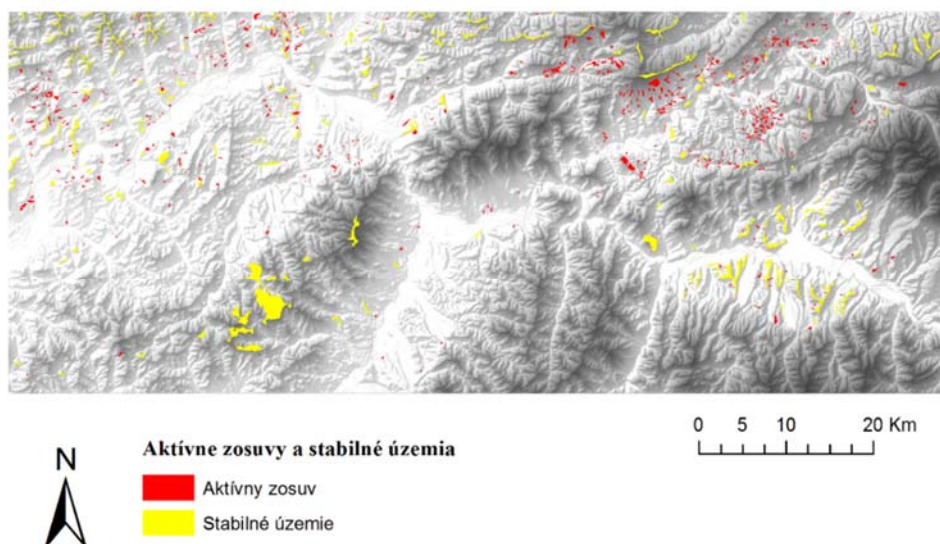
- Geologická mapa SR 1:50 000,
- Mapa zosuvov a stabilných území SR 1:50 000,
- Krajinná pokrývka (z angl. *corine land cover*).

Z nich boli vytvorení štyri vstupné parametre, a to geologické pomery daného územia, sklon svahov (*Obr. 5*), orientácia svahov, krajinná pokrývka. Tieto parametre bolo potrebné najprv vhodne pripraviť. Príprava dát bola vykonaná v softvéri ArcMap 10.2.

Zosuvné a stabilné oblasti boli náhodne rozdelené v pomere 80:20, pre potreby tréovania umelej neurónovej siete. Vzorka obsahujúca 20 % zosuvných a stabilných oblastí bude v budúcnosti využitá na validáciu tohto modelu. Takto boli vytvorené masky (*Obr. 6*) pomocou, ktorých sa neskôr extrahovali údaje z ostatných vrstiev. V tréovacej vzorke sú použité len aktívne zosuvné oblasti, ktorých bolo 442 a stabilných území 980. Po extrahovaní daných lokalít boli výsledné rastre exportované do ASCII súboru.



Obr. 5 Vypočítaný sklon reliéfu na celom modelovom území.



Obr. 6 Oblasti aktívnych zosuvov a stabilné územia

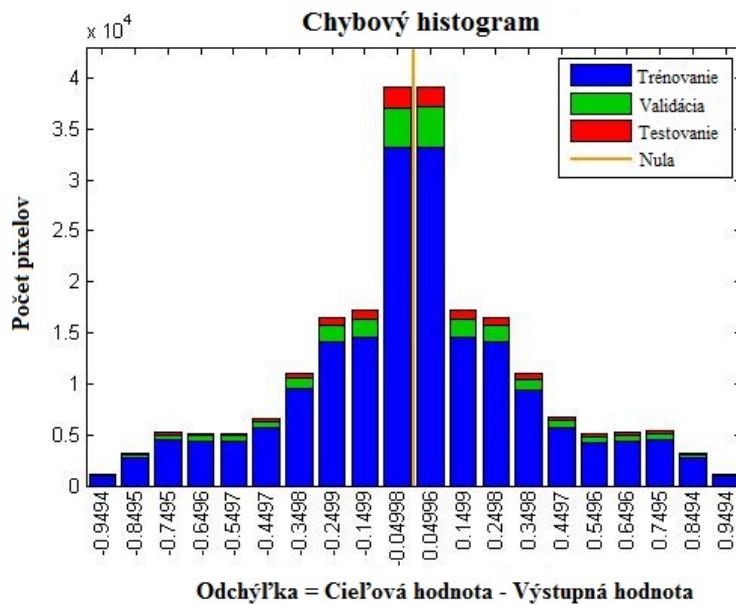
2.2 Tvorba trénovacej a cieľovej matice

Po vytvorení štyroch ASCII súborov bolo možné pristúpiť k vytvoreniu samotnej umelej neurónovej siete. Celá tvorbu umelej neurónovej siete prebehla v softvéri Matlab 13. Po načítaní jednotlivých súborov bola vytvorená trénovacia a cieľová matica hodnôt pre neurónovú sieť. Trénovacia matica obsahovala 4 riadky a 110 471 stĺpcov. Štyri riadky predstavovali 4 parametre, podľa ktorých sa mala sieť natréňovať a 110 471 stĺpcov predstavovalo počet jednotlivých pixelov, prislúchajúcich už spomínaným rastrom. Cieľová matica obsahovala dva riadky a 110 471 stĺpcov. Spomínaná matica bola tvorená hodnotami 1 a 0. Hodnota 1 prislúchala pixelu, ktorý predstavoval zosuvné územie a hodnota 0 zodpovedala pixelu, ktorý predstavoval stabilné územie.

2.3 Trénovanie siete

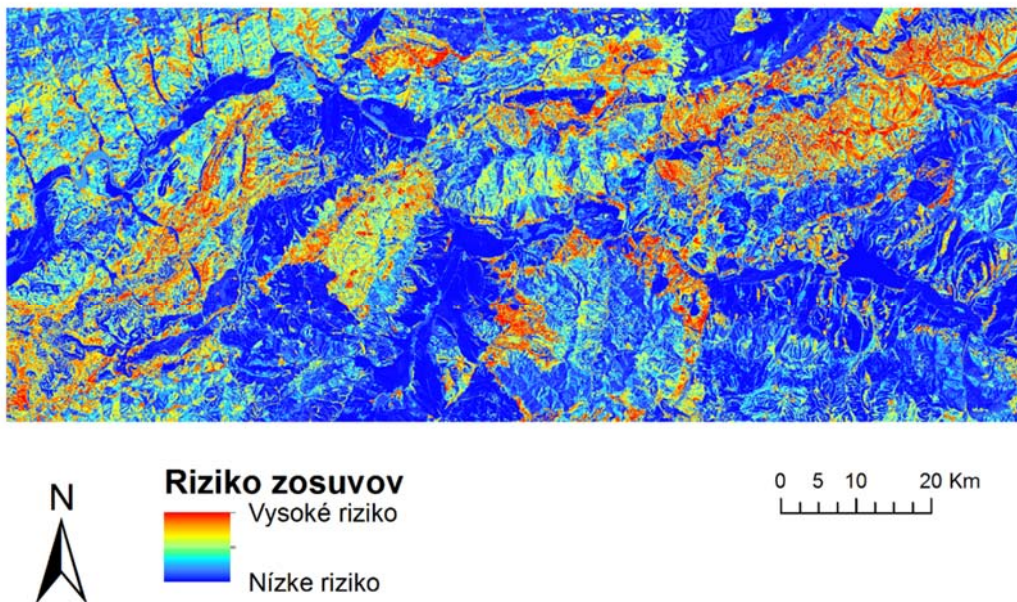
Po vytvorení trénovacej a cieľovej matice sme mohli začať s trénovaním (učením) umelej neurónovej siete. Na trénovanie boli využité už spomínané algoritmy GDM a SCG.

Voľba počtu skrytých neurónov je veľmi dôležitým prvkom v procese tvorby umelej neurónovej siete. V prípade výberu príliš veľkého počtu skrytých neurónov, môže dôjsť k tzv. preučeniu neurónovej siete. Naopak, ak by počet skrytých neurónov bol malý, neurónová sieť by nemusela správne klasifikovať hodnoty. Otázka výberu počtu skrytých neurónov je vo všeobecnosti často riešená náhodným skúšaním počtu skrytých neurónov. Niektorí autori uvádzajú dva spôsoby voľby najvhodnejšieho počtu skrytých neurónov. V prvom spôsobe sa začína na veľmi malom počte skrytých neurónov a postupne sa ich počet zvyšuje. Pri trénovaní sa sleduje priebeh výpočtu a počet správne klasifikovaných (naučených) hodnôt. Počet skrytých neurónov sa zvyšuje až dovtedy, kým sledované ukazovatele nezačnú vykazovať horšie hodnoty. Druhý spôsob je presne opačný. Začne sa s menším počtom neurónov a postupne sa ich počet postupne zvyšuje. Algoritmus vykazoval najlepšie výsledky pri 20 skrytých neurónoch. Toto tvrdenie je reprezentované histogramom chýb (Obr. 7).



Obr. 6 Chybový histogram predstavujúci rozdiel medzi cieľovými a výstupnými hodnotami

Validáciu a testovanie zobrazené v histograme vykonal softvér. Táto validácia výsledkov nie je postačujúca na vyjadrenie, že zvolený počet skrytých neurónov je najvýhodnejší. Po natrénovaní umelá neurónová sieť klasifikovala celé modelové územie. Po klasifikácii bola pomocou vstavaných príkazov exportovaná výsledná matica do ASCII súboru. Výsledná matica obsahovala hodnoty pravdepodobnosti výskytu zosuvu resp. absencie zosuvu pre daný pixel. ASCII súbor následne konvertoval do rastrovej mapy (Obr. 8) a bol vhodne graficky upravený.

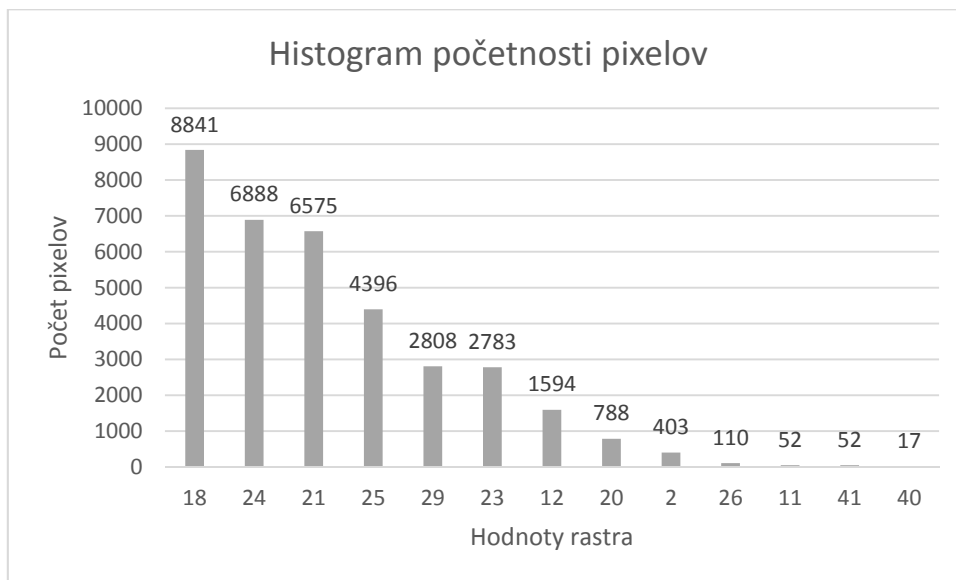


Obr. 7 Výsledná rastrová mapa rizika zosuvov z neurónovej siete s 20 skrytými neurónmi vytvorená metódou SCG

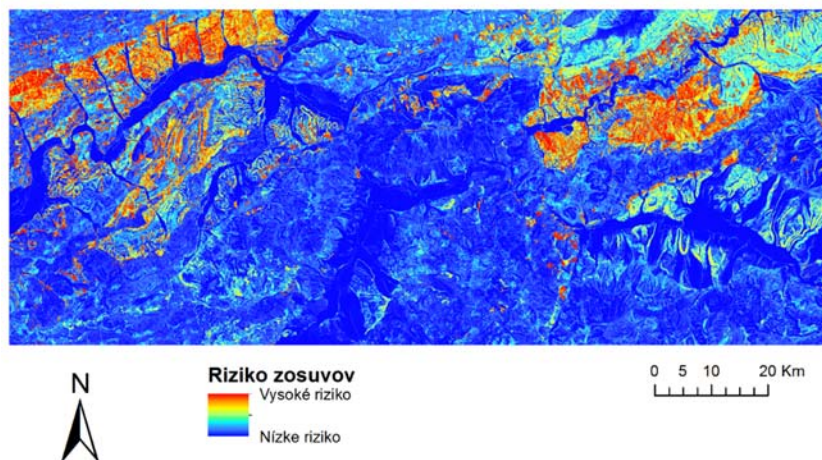
2.4 Reklasifikácia

Pri tréovaní siete vznikla domnienka, ako sa niektoré hodnoty vstupných parametrov podieľajú na samotnom výpočte. Konkrétne sa jednalo o hodnoty geologickej skladby podłożia a krajinej pokrývky. Geologická mapa SR 1:50 000 ako aj krajinná pokrývka obsahujú číselné hodnoty v intervaloch od 1 až po 241 pre geologickú mapu a od 1 až po 41 pre krajinnú pokrývku. Tieto hodnoty prislúchajú jednotlivým geologickým vrstvám resp. jednotlivým typom pôdneho krytu. Otázkou bolo ako sa tieto hodnoty podieľajú na učení umelej neurónovej siete a jej následnej interpretácii naučených poznatkov.

Pre tento účel boli vytvorené histogramy (Obr. 9) početnosti jednotlivých tried. Jednotlivé čísla reprezentujú číselné označenie vrstiev a nie ich názov. Každý jeden stĺpec predstavuje počet pixelov daného rastra, ktoré prislúchajú hodnote zodpovedajúcej územi, na ktorom sa vyskytol zosuv. Čiže z týchto histogramov je jasné, aké geologické vrstvy resp. krajinné pokrývy sú náchylné na svahové deformácie.



Obr. 8 Histogram početnosti pixelov krajinej pokrývky



Obr. 9 Výsledná rastrová mapa rizika zosuvov z neurónovej siete s 20 skrytými neurónmi vytvorená metódou SCG po reklasifikácii

Po vytvorení histogramov bolo možné pristúpiť ku reklasifikácii rastrov. Z reklasifikovaných rastrov boli opäť extrahované hodnoty zo zosuvných území a exportované do ASCII súboru. Rovnako ako pre nereklasifikované údaje bola vytvorená umelá neurónová sieť a z nej vznikol predikčný model zosuvov. Na tvorbu umelej neurónovej siete boli opätovne použité algoritmy SCG a GDM. Na obrázku 10 je uvedený spomínaný predikčný model.

Podrobný postup tvorby modelu je uvedený v bakalárskej práci [6].

Záver

Z uvedených predikčných modelov jasne vyplýva potreba ďalšieho výskumu a zaoberania sa o umelé neurónové siete. Modely ukazujú, že územie zvolené v tejto práci má množstvo potenciálne nebezpečných zosuvných zón. Avšak jednoznačné závery nemožno vyvodiť, pretože tieto modely neboli pre krátkosť času, náročnosť a komplexnosť danej problematiky validované. Predmetom ďalšieho skúmania by mala byť validácia dosiahnutých výsledkov, testovanie ďalších parametrov siete, ale aj pridanie ďalších faktorov ovplyvňujúcich vznik zosuvov.

Literatúra

- [1] NEMČOK, PAŠEK, RYBÁŘ.: Delenie svahových deformácií. Praha: Zbor. Geol. Vied, r.HIV, 1974.
- [2] SINČÁK P. – ANDREJKOVÁ G.: Neurónové siete Inžiniersky prístup (1.diel)
- [3] KVASNIČKA V. a kol.: Úvod do neurónových sietí. Bratislava. IRIS 1997, ISBN 80-88778-30-1.
- [4] PETRÝDESOVÁ L., LIŠČÁK P.. Monitorovanie vybraných parametrov podzemných vôd na Veľkomarskom zosuve (Liptovská kotlina). In ACTA GEOLOGICA SLOVACA. 2009, roč 1, č.2, s .83-91
- [5] TORNAYI R, DUNČKO M.: Použitie bivariačnej a multivariačnej analýzy na hodnotenie zosuvného hazardu Kysuckého regiónu (Západné Karpaty). In ACTA GEOLOGICA SLOVACA. 2013, roč 5, č.2, s. 179-193
- [6] JURČIAK F. (2014). Aplikácia umelých neurónových sietí na tvorbu predikčných modelov v prostredí GIS. Bakalárska práca, Slovenská technická univerzita v Bratislave, Stavebná fakulta, Bratislava, 2014.