

NEURÓNOVÉ SIETE

PROJEKT 2

SELF-ORGANIZING MAP

Autor: Marián Kravec

Úvod

V tejto úlohe sa snažíme natrénovať dvojrozmernú štvoruholníkovú SOM na vizualizáciu 8-rozmerných dát (z toho jedna kategória). Ide o dataset seeds z UCI Machine Learning Repository.

Dáta

Máme dataset tvorený 210 dátovými bodmi ktoré majú 8 rozmerov, 7 rozmerov sú parametre bodu a ôsmi je kategória. Tento dataset rozdelíme v pomere 5:2 na tréningové a testovacie dáta. Takto získame 150 bodov na tréningovanie a 60 bodov na testovanie, pričom v oboch datasetoch bolo približne rovnaké zastúpenie všetkých tried (trénovací: {1 : 51, 2 : 52, 3 : 47}, testovací: {1 : 19, 2 : 18, 3 : 23}).

Architektúra a hyperparametre

Pri výbere modelu sme skúšali tri rôzne normy pre vzdialenosti na sieti, konkrétne sme skúsili normy L_1 , L_2 , L_{max} . Zároveň sme skúsili trénovať s diskretnou aj spojitou funkciou susednosti. V neposlednej rade sme skúšali aj viac štartovacích hodnôt parametra α , konkrétne sme skúsili hodnoty: {0.5, 0.7, 1, 2, 5, 10} (finálna hodnota α bola pre všetky modely 0.01)

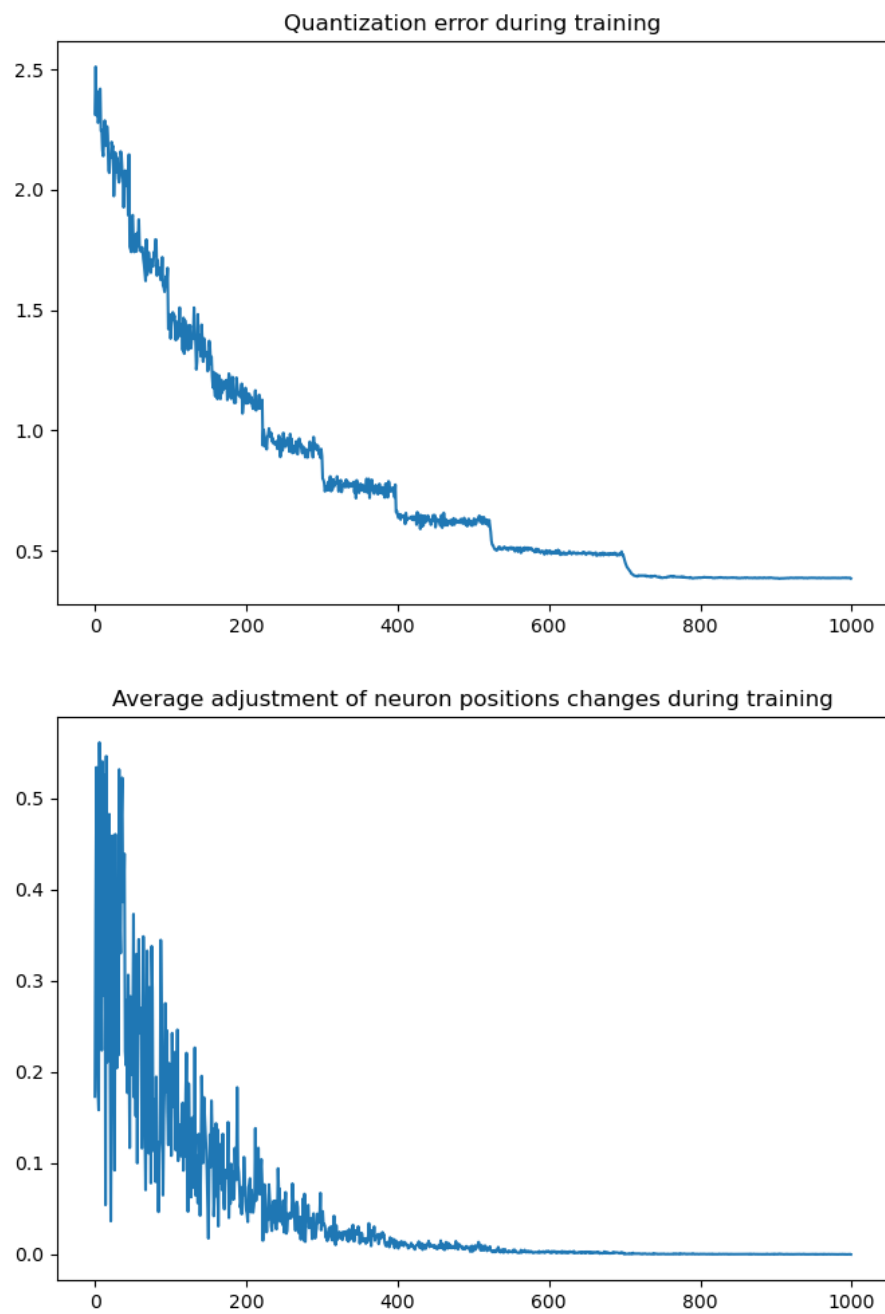
Všetky modely boli trénované na 500 epoch. Veľkosti všetkých trénovaných sietí boli 10×10 . Parameter λ začína na hodnote priemeru rozmerov siete ($\frac{\#rows + \#columns}{2}$) aby sa na začiatku zmena propagovala po celej sieti a končí na hodnote 1 aby ku koncu bola zmena iba lokálna.

Nakoniec sa ako najlepší model ukázal model využívajúci normu L_1 , diskretnú funkciu susednosti a počiatočný parameter $\alpha = 1$.

Výsledky modelu

Výsledný model sme trénovali na 1000 epoch s parametrami najlepšieho modelu z predchádzajúcej časti.

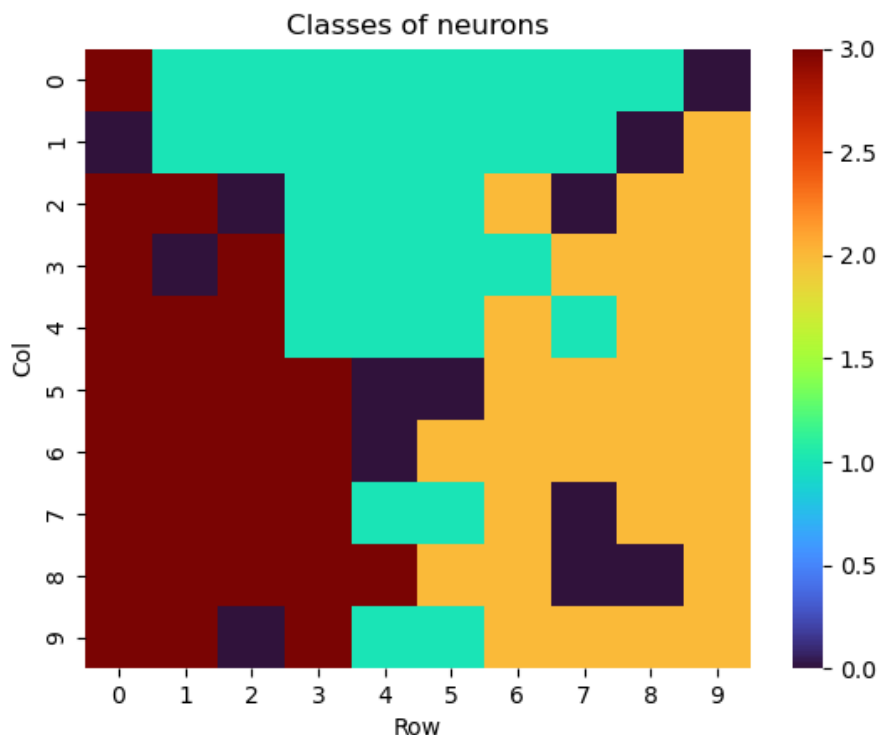
Ak si vykreslíme kvantizačnú chybu a priemernú zmenu pozície neurónov počas tréningovanie dostaneme takéto grafy:



Obr. 1: Quantizačnej chyby a priemernej zmeny pozície neurónu počas trénovania modelu

Vidíme, že quantizačná chyba klesá pomerne skokovo keď určitý počet epoch sa chyba mení minimálne a náhle v jednej epoche klesne výraznejšie. Finálna chyba je približne 0.375. V prípade priemernej zmeny pozície neutrónov vidíme, že postupne klesá a po približne 600 epochách je iba minimálna.

Keď si pozrieme ktoré neuróny aktivujú jednotlivé triedy dostaneme takýto výsledok:



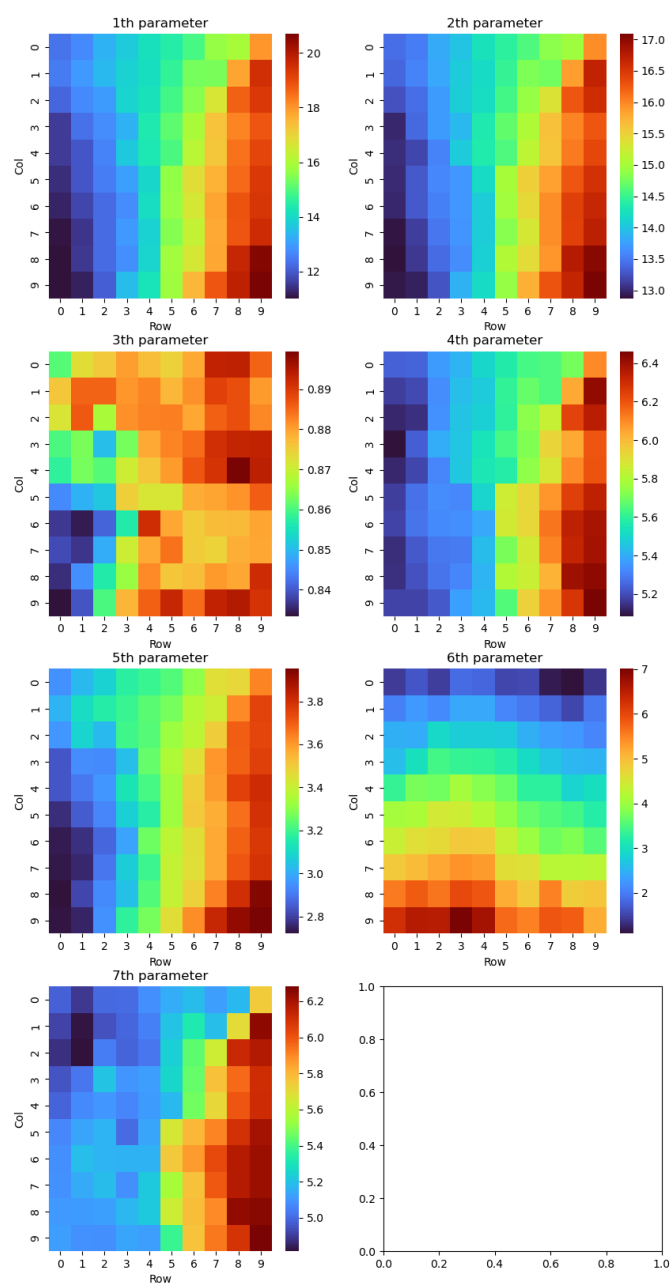
Obr. 2: Triedy aktivujúce jednotlivé neuróny

Vidíme, že na hraniciach medzi triedami ostalo niekoľko neurónov ktoré nie sú aktivované žiadnou triedou ale inak vidíme pomerne pekné hranice jednotlivých tried (výnimkou je pár neurónov klasifikovaných ako trieda 1 na hranici tried 2 a 3).

Na základe tohto výsledku sa môžeme pokúsiť klasifikovať našu testovaciu množinu vstupov. Pri porovnaní predikcii modelu a skutočných tried sme zistili, že náš model klasifikoval správne 81.7% testovacích bodov. Pričom rozdelenie tried správne klasifikovaných bodov bolo nasledovné: {1 : 13, 2 : 14, 3 : 22}. Čiže najhoršie výsledky dosiahol pre triedu 1 kde správne klasifikoval iba približne 68% bodov, o niečo lepšia bola trieda 2 kde správne klasifikoval takmer 78% a úplne najlepšou bola trieda 3 kde bolo správne klasifikovaných až vyše 95% bodov. Ak si pozrieme celú confusion matrix zistíme, že veľká časť chýb je spôsobená tým, že testovací vstup skončí v mŕtvom neuróne:

očakávanie/realita	1	2	3	0
1	68.42%	0.0%	15.79%	15.79%
2	5.56%	77.78%	0.0%	16.67%
3	0.0%	0.0%	95.65%	4.35%

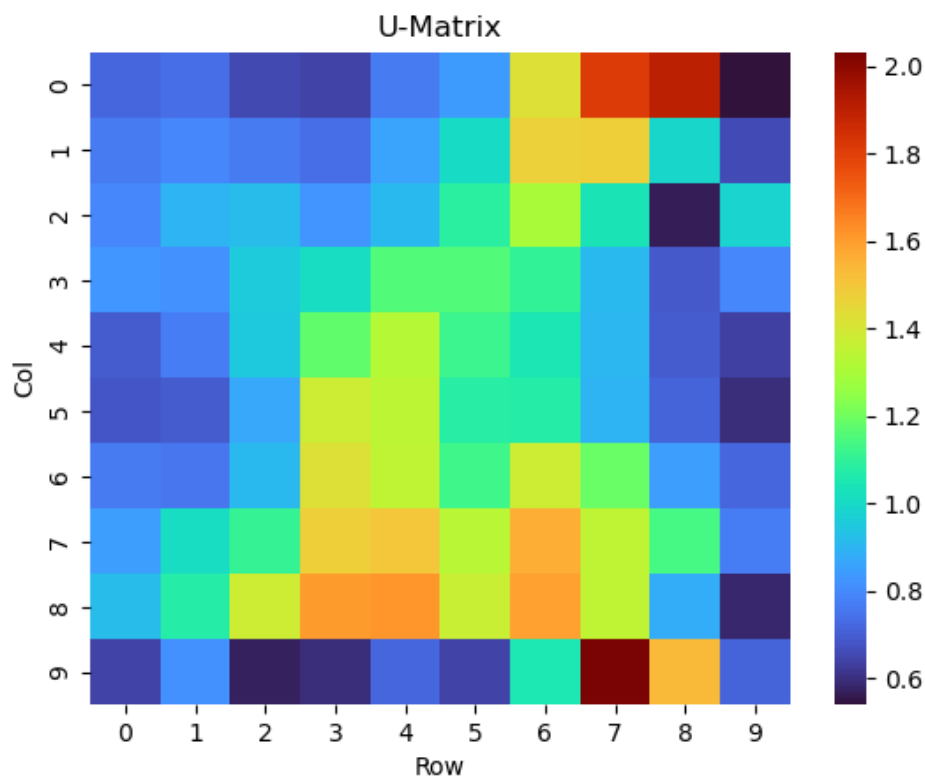
Ako ďalšie si môžeme pozrieť heatmapy pre jednotlivé parametre modelu (máme ich 7):



Obr. 3: Heatmapy jednotlivých parametrov modelu

Na týchto heatmapách si môžeme všimnúť, že pre parametre 1, 2 a 5 vyzerajú takmer totožne a pre parametre 4 a 7 sú veľmi podobné pričom podobnosť je aj s rozdelením tried (na grafe vyššie). Heatmapy pre parametre 3 a 6 sú diametrálne odlišné od ostatných aj navzájom.

Na záver si ešte vizualizujeme U-matrix pre našu SOM:



Obr. 4: U-matrix

Úprimne, si nie som istý ako tento výsledok interpretovať, teoreticky by mal znamenať, že model našiel 2 (teoreticky 3) clustre dát, jeden vidíme hore v pravom rohu a druhý dole v strede matice (teoreticky tretí dole v pravom rohu). Avšak tento výsledok je výrazne odlišný od výsledku tried pre jednotlivé neuróny, čo by mohlo indikovať zle natrénovanú sieť, avšak v takom prípade by sme nepredpokladali tak vysokú presnosť zaradzovania do tried ($> 80\%$).