

# 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.

## Architektúra a hyperparametre

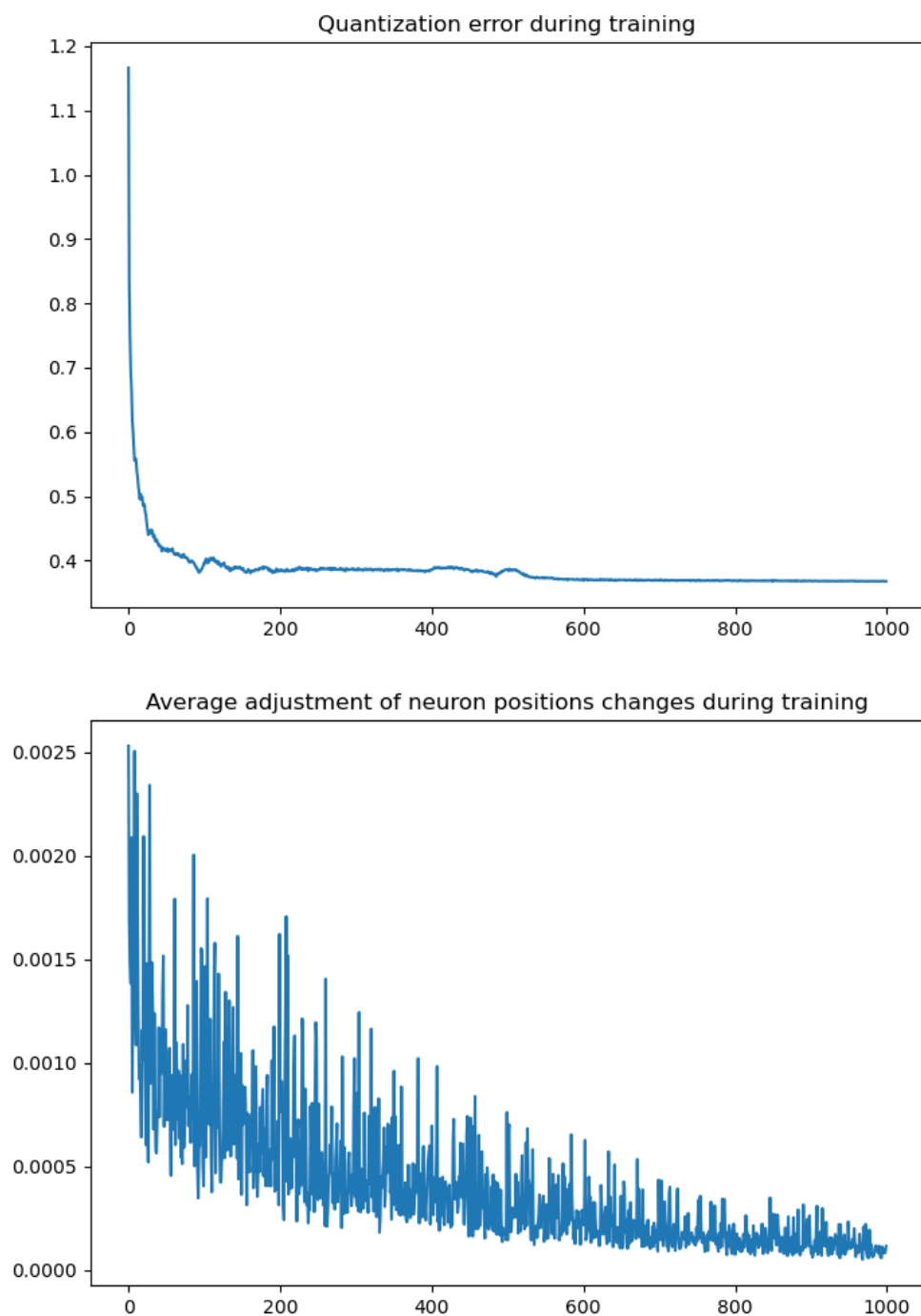
Pri výbere modelu sme skúšali tri rôzne normy pre vzdialenosti na sieti, konkrétne sme skúsili normy  $L_1$ ,  $L_2$ ,  $L_{max}$ . Zároveň sme skúsili trénovať s diskretnou aj spojitou funkciou susednosti. V neposlednej rade sme skúšali aj viac štartovacích hodnôt parametra  $\alpha$ , konkrétne sme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

Všetky modely boli trénované na 500 epoch. Veľkosti všetkých trénovaných sietí boli  $10 \times 10$ . Parameter  $\lambda$  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  $\alpha = 0.7$ .

## Výsledky modelu

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



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