

Generovanie váh pre CNN pomocou symbolickej regresie a genetického programovania

Bc. Richard Seipel

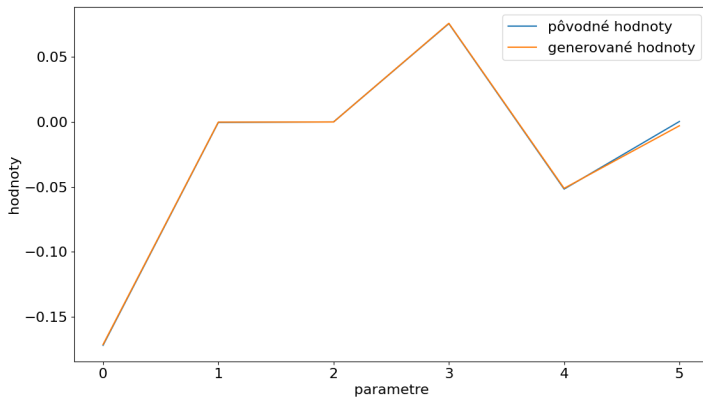
Fakulta informačních technologií Vysokého učení technického v Brně
Božetěchova 1/2. 612 66 Brno - Královo Pole
xseipe00@fit.vutbr.cz



- Konvolučná neurónová sieť typu **LeNet**.
 - 2 konvolučné vrstvy (156 a 2416 parametrov).
 - 3 plne prepojené vrstvy (30840, 10164 a 850 parametrov).
- Dátová sada MNIST.
- Referenčná presnosť siete 92%.

- Priama **multidimenzionálna** regresia.
- Pôvodný počet rozmerov.
- Vysoká komplexita → slabé výsledky.

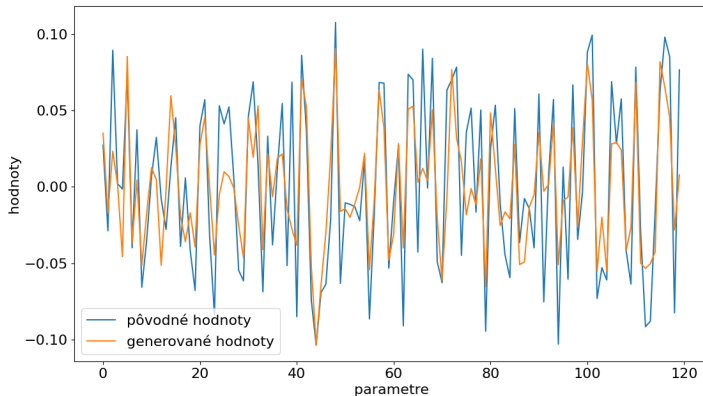
- Inšpirácia Fourierovým radom pri voľbe funkcií.
- Vrstvy transformované do **jedného rozmeru**.
- Uspokojivé výsledky pri menších vrstvách .
 - Prvá konvolučná vrstva - **90%**.
- Nedostatočné výsledky pri vrstvách s viacerými parametrami (konštantná funkcia, sínusoida).



Obrázek: Aproximácia bias parametrov prvej konvolučnej vrstvy.

- Cieľom nájsť *kompromis* v počte a tvare rozmerov.
- Vhodné rozdelenie veľkých vrstiev na *menšie skupiny*.
- Zvolené skupiny:
 - Konvolučné vrstvy rozdelené podľa filtrov → v jednej skupine všetky **konvolučné jadrá jedného filtra**.
 - Plne prepojené vrstvy rozdelené po neurónoch → v jednej skupine všetky **vstupné váhy pre jeden neurón** (dlhší vektor).
- Zvolené rozmery:
 - Konvolučné vrstvy → *počet filtrov* x (**počet kanálov** x **výška**) x **šírka konvolučného jadra**.
 - Plne prepojené vrstvy → počet výstupných neurónov x **počet vstupných neurónov**.

- Výsledky:
 - Dosiahnutie rozumnej *presnosti pre väčšie vrstvy*.
 - Prvá konvolučná vrstva - 91%.
 - Druhá konvolučná vrstva - 80%.
 - Druhá plne prepojená vrstva - 77%.
 - Tretia plne prepojená vrstva - 74%.
 - Nevýhodou vyšší počet behov a získaných funkcií (16 a 84).



Obrázek: Aproximácia parametrov jedného vektora druhej plne prepojenej vrstvy.

- Snaha o *zjednodušenie a zmenšenie matíc* parametrov.
- Nie je potreba rozdeľovať na skupiny → stačí *menší počet výrazov* popisujúcich parametre.
- Potreba navyiac ukladať maticu vektorov komponent.
- Pôsobivé, no nestabilné výsledky.
 - Napr. druhá konvolučná vrstva bez delenia na skupiny - 84%.
Pre 16 komponent matice 16×16 (256) prvkov pre $16 \times 6 \times 5 \times 5$ (2400) parametrov.

- Experimentovanie s rôznymi prístupmi.
- Najstabilnejšiou metóda rozdelenia na skupiny.
- Pre lepšie výsledky potrebných viac experimentov.
- Rozsiahlejšie experimenty s parametrami symbolickej regresie → možné zmenšenie dĺžky výstupov regresie a zvýšenie presnosti.