

Domáca úloha č. 4

2-AIN-150, Zima 2020

Termín: 27.11.2020, 23:59, boza@fmph.uniba.sk, predmet mailu: Strojove ucenie DU 4

Skôr ako sa pustíte do riešenia domácej úlohy, oboznámte sa so všeobecnými pokynmi, ktoré sú priložené na konci tohto dokumentu. Riešenia, ktoré odovzdáte, musia byť vaše vlastné. Neopisujte a nesnažte sa nájsť riešenia v literatúre alebo na internete!

Optimalizácia hyperparametrov

Poriadne čítajte, čo treba odovzdať.

Uvažujme mierne pozmenenú neurónovú sieť z tretích cvičení nad datasetom Fashion MNIST.

Naša sieť má niekoľko skrytých vrstiev medzi ktorými je ešte dropout vrstva. Dropout vrstva počas tréningu náhodne zahadzuje neuróny, čo má za následok menší overfitting. Množstvo zahodených neurónov sa dá nastaviť hyperparametrom.

Túto sieť ale treba nastaviť. Budeme nastavovať: počet a veľkosť skrytých vrstiev (musia byť rovnako veľké), silu dropoutu a počet tréningových epóch. Chceme vybrať nastavenia, ktoré minimalizujú validačnú chybu.

Vašou úlohou je implementovať a porovnať dva prístupy na nastavovanie hyperparametrov.

Prvý prístup sa nazýva random search. Ako názov vypovedá skúša náhodne nastavenia parametrov (samozrejme veľa krát) a vyberie najlepšie z nich.

Druhý prístup sa volá grid search. Pre každý parameter si zvolí niekoľko možných hodnôt (niekde to je lineárna škála, niekde logaritmická) a potom vyskúša všetky kombinácie. V niektorých prípadoch ešte okolo najlepšej kombinácie robí ďalšiu iteráciu grid searchu (akoby prizožmoval), ale toto implementovať nemusíte.

Implementujte obidva prístupy a porovnajte ich. Ktorý prístup vám dáva rýchlejšie dostatočne dobrý výsledok?

Odovzdajte notebook z cvičení, kde na konci doprogramujete random search a grid search nad popísanou architektúrou siete. Samozrejme nazabudnite ich aj pustiť. Takisto pridajte krátky komentár (a prípadne vhodné grafy) o tom, ktorá metóda sa vám zdá byť lepšia. Pozn.: Máte zakázané volať knižničné funkcie, ktoré grid search a random search implementujú za vás.

Takisto máte k dispozícii súbor `architecture.py` s architektúrou siete a rozsahom parametrov, ktoré stačí uvažovať.