

POSUDEK DIPLOMOVÉ PRÁCE (POSUDEK ŠKOLITELE).

Student: Jan Tichavský

Název práce: Trénování neuronových sítí založené na vysoce robustních odhadech

Předložená práce se věnuje vícevrstvým perceptronům, tedy velmi rozšířené skupině umělých neuronových sítí, a jejich senzitivitě (nerobustnosti) vzhledem k přítomnosti odlehlých měření v datech při řešení úlohy nelineární regrese. Klade si za cíl navrhnout nové vhodnější (robustnější) postupy pro jejich trénování a ilustrovat je na numerických příkladech.

První myšlenkou je robustně modelovat spojitou odezvu pouze na základě pozorování, které leží mezi dvěma nelineárními kvantily, přičemž oba jsou odhadnuty pomocí vícevrstvého perceptronu. Druhou myšlenkou je nahradit ztrátovou funkci ve vícevrstvěm perceptronu takovou robustní verzí, která se používá pro vysoce robustní odhady v lineární regresi. Autor popsal oba postupy, implementoval je v prostředí Python a ilustroval jejich chování na simulovaných i reálných datech.

Práce je rozdělena do 7 kapitol, známé poznatky prezentuje v kapitolách 1-3, kde nelineární regresní úloze je věnována kapitola 1 a neuronovým sítím kapitola 2. Nové výsledky v dané oblasti jsou shrnuty v samostatné kapitole 3. Nové metody jsou popsány v kapitole 4 a výsledky výpočtů v kapitole 5. Kapitola 6 přidává i porovnání metod pomocí neparametrického bootstrapu. Kapitola 7 obsahuje popis souborů, které jsou obsaženy na přiloženém CD.

K textu práce bych měl tyto konkrétní připomínky (některé jsou drobnosti).

- V kapitolách 1 a 2 autor píše o perceptronu, ale evidentně má na mysli vícevrstvý perceptron (tj. model neuronové sítě, zatímco perceptronem se rozumí jen model s jediným neuronem). Od kapitoly 3 je pak terminologie už správná.
- Str. 13. TMSE nezanedbá čtvrtinu největších odchylek, ale čtvrtinu největších *absolutních* odchylek. Stejně i na str. 40 a 41.
- Str. 13, definice kvantilů. Obvykle se kvantily definují jinak, v práci jde spíš o jejich vysvětlení než přesné definice, kterých existuje několik typů. Například software R udává dvacetiprocentní kvantil dané množiny M jako 4,6.
- Str. 14. Není jasné, zda normalita náhodných regresních chyb e představuje předpoklad; v lineární regresi se obvykle tato normalita nepotřebuje, pokud se jen odhaduje β (a neprovádí například testování hypotéz).
- Vzhledem k názvu práce bych čekal vysvětlení, co se rozumí *vysoce* robustními odhady.
- Nejednotné značení napříč celou prací: f značí na str. 9 odhadnutou funkci (odhad F), na str. 20 aktivační funkci, na str. 35 náhodnou veličinu v simulacích.

- Str. 31 uvádí, že se architektura hledá jako optimální (patrně přes různé avšak pevně zvolené hodnoty jednotlivých parametrů). V příkladech kapitoly 5 není jasné, zda se architektura opravdu optimalizuje, resp. není uvedeno přes jaké různé pevně zvolené hodnoty jednotlivých parametrů.
- Str. 35 uvádí, že je třeba pro výpočet zvolit B (*batch size*). Není jasné, jaká volba byla využita v příkladech kapitoly 5 (s výjimkou příkladu na str. 42, kde je explicitně uvedena volba $B = 100$).
- Příklad v kapitole 6 je zajímavý tím, že IQ-MLP má nejužší konfidenční pás kolem odhadnutého trendu, ale nevíme, nakolik je samotný IQ-MLP odhad dobrý; jeho predikční chyba není vyčíslena a porovnána s jinými metodami.

Z formálního hlediska odpovídá počet překlepů rozsahu práce. Z některých drobností lze uvést, že některé odstavce jsou v textu odsazené, některé ne; správný pravopis by měl být např. TensorBoard (str. 25) nebo CRC Press (str. 53); česky se říká *součín* na místo *produkt* (str. 27). Zejména v popisu složitějších metod se autor snaží vysvětlit postupy svými slovy, neopisuje z literatury, i za cenu toho, že jazyk pak nemusí odpovídat odbornému textu.

Student splnil zadání své diplomové práce. Příklady ukazují, že nově navržené robustní postupy dávají smysl a jsou výrazně vhodnější pro kontaminovaná data než obvykle používané neuronové sítě. Celkově doporučuji ohodnotit diplomovou práci známkou **B**.

Případné dotazy k obhajobě:

- Jak bylo při výpočtech zvoleno η , definované na str. 25?
- Co znamená *účinný* odhad (str. 27)?

RNDr. Jan Kalina, Ph.D.
Ústav informatiky AV ČR, v.v.i.

V Praze, dne 8. července 2020