

## **Posudek na disertační práci**

Mgr. Terezie Šámalová: "Learning from Data based on Approximation of Functions by Neural Networks"

Disertační práce mgr. T. Šámalové se zabývá analýzou aproximačních schopností dopředné neuronové sítě s jednou skrytou vrstvou a souvisejících prakticky použitelných algoritmů.

Téma aproximačních schopností (zejména přesnost aproximace) neuronových sítí je studováno dlouhodobě od vzniku prvních paradigmat neuronových sítí. Vlastnosti a tedy i chování neuronových sítí vycházely z experimentálních pozorování a studium exaktních vlastností se rozvinulo později až s rozšířením neuronových sítí jako nástrojem pro zpracování dat. Z toho pohledu je práce aktuální, protože hledá a popisuje aproximační schopnosti neuronových sítí (v případě předložené práce zaměřené úžeji na dopřednou neuronové sítě s jednou skrytou vrstvou) s pomocí exaktního matematického aparátu. Z pohledu zaměření předložené práce, která se orientuje na aproximační vlastnosti dopředné neuronové sítě s jednou skrytou vrstvou navazuje práce v mnoha ohledech na práce dr. Kůrkové a jejích kolegů.

Úvodní kapitola stručně a výstižně uvádí pojmy z oblasti neuronových sítí a teorie aproximace a jejich základní vlastnosti, o které se opírá další text práce. Vlastní výsledky práce jsou soustředěny zejména do 2. kapitoly zaměřené na odhad rychlosti aproximace (*rates of approximation*<sup>1</sup>, str. 12), do 3. kapitoly zaměřené na iterativní a pravděpodobnostní algoritmy pro odhad aproximačních schopností algoritmů, zejména neuronových sítí a částečně do 4. kapitoly navazující na předchozí kapitoly a zaměřené na prořezávání rozsáhlých sítí.

Ve 2. kapitole zaměřené na odhad rychlosti aproximace neuronovými sítěmi se autorka zabývá základní otázkou jak přesně lze aproximovat danou funkci omezeným počtem neuronů. Vychází zde z průkopnických prací Maureyho, Jonese, Barrona a Darkena a kol. a zejména pak z práce Kůrkové, Kainena a Kreinoviche, na kterou úzce navazuje a rozšiřuje a zobecňuje výsledky této práce. Tato vylepšení umožňují přímější a obecnější aplikaci výsledků zmíněných průkopnických prací výše uvedených autorů. V práci jsou rozvedeny zejména vlastnosti týkající se tzv. G-variace (G-variation) zavedené dr. Kůrkovou. Pro odhad G-variace je zapotřebí reprezentovat funkci v integrálním tvaru. Pro tento účel jsou v práci uvedeny známé příklady takových reprezentací včetně autorkou prokázané možné speciální integrální reprezentace funkce v Sobolevově prostoru. V závěru kapitoly autorka uvádí několik přímočarých důsledků vět týkajících se teoretických mezí rychlosti aproximace uvedených v předchozích částech kapitoly. Výhradu mám zejména k aktuálnosti těchto výsledků. Uvedená řada vět a důsledků týkajících se G-variace navazuje na cca 10 let staré práce, převážně z roku 1997. Předložené výsledky považuji za korektní (i když místy ne zcela přesně formulované), současně ale nejsou příliš průkopnické.

---

<sup>1</sup> Zde pojem *rates of approximation* překládám do češtiny poněkud nepřesně jako *rychlost aproximace*, vhodnější je původní anglický název.

Ve 3. kapitole se autorka zaměřuje na způsoby odhadu rychlosti aproximace. Jedna cesta vede přes iterativní zlepšování odhadu na základě předchozích aproximujících funkcí, což je cesta, která byla položena Jonesem a Barronem. Druhý přístup je založen na pravděpodobnostních metodách, jehož základy dal Maurey a na který navazuje i autorka a analyzuje vlastnosti tohoto pravděpodobnostního přístupu ve srovnání s iterativními metodami. Jak autorka ukazuje ve své práci, pravděpodobnostní algoritmy jsou vhodné pro případy rozsáhlých sítí neuronů, kde jsou využity pro prořezávání těchto sítí s cílem zlepšit aproximační vlastnosti. Kapitola teoreticky rozebírá různé algoritmy a jejich vlastnosti s ohledem na jejich aproximační schopnosti a časovou náročnost. Závěr kapitoly je věnován srovnání obou přístupů, iterativního a pravděpodobnostního. Časové složitosti algoritmů nicméně silně závisí na řešené úloze, takže dílčí závěry lze činit jen za jistých omezujících předpokladů. Výhodou pravděpodobnostních algoritmů je bezesporu skutečnost, že mohou být využity pro jiná paradigmatu neuronových sítí, jako jsou např. RFB sítě. Troufám si tvrdit, že hlubší rozpracování tématu této kapitoly by mohlo přinést zajímavější výsledky než jsou výsledky uvedené v 2. kapitole.

4. kapitola se zabývá prořezáváním nalezeného řešení (aproximace funkce neuronovou sítí) s využitím pravděpodobnostního přístupu. Využit je přitom již uvedený algoritmus ze 3. kapitoly – Algoritmus 3.2.2. Hlavní přínos kapitoly pak lze spatřovat ve větě 4.3.1, která zajišťuje, že získané výsledné řešení (neuronová síť) výrazně neovlivní kvalitu původní aproximované funkce.

Zásadní nedostatek práce spatřuji ve skutečnosti, že v uvedeném seznamu literatury se nevyskytují významnější práce autorky. Očekával jsem zejména několik prací publikovaných ve vědeckých časopisech. Místo toho jsou uvedeny pouze podružné práce, vesměs nerecenzované, jako jsou např. technické zprávy (3) nebo sborníky z Doktorandského dne (4). Nenašel jsem ani společnou publikaci se školitelkou, dr. Kůrkovou.

Očekával bych také, že v práci bude uvedeno více příkladů, které budou demonstrovat užitečnost vyslovených (a dokázaných) vět.

Z předložené práce je zřejmé, že autorka má znalosti problematiky aproximace funkcí a neuronových sítí a že se v dané problematice orientuje. Vlastní přínos práce lze spatřovat v řadě zformulovaných vět a důsledků z oblasti aproximace funkcí a pravděpodobnostních algoritmů, které posouvají znalosti v dané oblasti, přičemž za nejužitečnější považuji výsledky v 2. kapitole.

Po formální stránce nemá práce nedostatky, snad jen zvyk, že se soupis literatury dává před přílohy a ne za ně.

Práce splňuje podmínky samostatné tvůrčí vědecké práce a obsahuje původní výsledky. Stanovený cíl práce považuji za splněný a práci **doporučuji** k obhajobě.

27. srpna 2008

