

# OPONENTSKÝ POSUDEK DISERTAČNÍ PRÁCE

**Téma práce:** Learning from Data based on Approximation of Functions  
by Neural Networks

**Autor práce:** Terezie Šámalová, MFF UK a ÚI AV ČR, Praha

---

## Aktuálnost tématu práce

V souvislosti s aktuálním vývojem dostupné výpočetní techniky (sto a více-jádrové procesory) nabývá na významu i studium vlastností masivně paralelních výpočetních modelů, mj. umělých neuronových sítí, genetických algoritmů a dalších. Předkládaná disertační práce Terezie Šámalové patří do zmíněné perspektivní oblasti výzkumu. Předmětem práce je analýza aproximačních schopností vrstevnatých neuronových sítí (s jednou skrytou vrstvou) a návrh nových postupů umožňujících snížit výpočetní složitost vytvářených modelů.

Disertantka ve své práci vycházela z významných výsledků, které v prestižních světových časopisech publikovala její školitelka RNDr. Věra Kůrková, DrSc. a další. Publikované důkazy jsou konstruktivní a lze je tedy využít např. k odhadu počtu skrytých neuronů potřebných k dosažení požadované přesnosti aproximace, resp. pro tzv. inkrementální učení. Alternativou k výše uvedenému přímému přístupu jsou tzv. pravděpodobnostní metody, které by zřejmě bylo možné úspěšně použít i při prořezávání skrytých neuronů. Tyto metody však zatím v literatuře nebyly podrobně zkoumány.

## Cíl práce

Cílem práce bylo analyzovat aproximační schopnosti dopředných neuronových sítí s jednou skrytou vrstvou. Standardní typ vrstevnatých neuronových sítí totiž představuje univerzální aproximační model. Vlastnosti tohoto modelu byly (s ohledem na hustotu C-prostorů v  $L_p$ -prostorech) většinou studovány v prostoru spojité funkce a charakter publikovaných výsledků, bohužel, neumožňoval jejich masivnější nasazení v praxi.

Aplikace z teoretického pohledu optimálních aproximačních metod navíc v mnoha případech není možná, a to vzhledem k příliš velkému množství volných parametrů použitých modelů. Na základě provedeného rozboru vlastností studovaného modelu proto měla uchazečka navrhnout nové efektivnější metody použitelné pro jeho učení. Teoretické vlastnosti navrhovaných algoritmů měla disertantka ověřit.

## Metody zpracování

V úvodní části práce se autorka zabývá popisem základního modelu dopředných neuronových sítí a studiu známých výsledků týkajících se aproximačních schopností dopředných sítí s jednou skrytou vrstvou. Při studiu aproximačních schopností popisovaného modelu se používají dva základní přístupy. První směr, který vychází z výsledků Jonese, Barrona a Darkena, v zásadě analyzuje iterativní proces (inkrementální) aproximace funkce.

Další část práce je věnována druhému přístupu, který je založený na poznatcích Maureyho, Makovoze a dalších. Tento přístup konstruktivní není a jeho hlavní idea vychází z nenulové pravděpodobnosti nalezení vhodné aproximace hledané funkce pomocí pravděpodobnostních metod. Tato myšlenka představuje také základ nově navržené metody pro učení a prořezávání zkoumaných dopředných sítí s jednou skrytou vrstvou.

## Dosažené výsledky

Těžiště práce spočívá ve studiu aproximačních schopností dopředných neuronových sítí s jednou skrytou vrstvou. Na základě poměrně rozsáhlé rešerše známých poznatků disertantka některé z uvedených výsledků zobecnila (pro funkce z  $L_p$ ). Mimo jiné to umožňuje např. lépe zohlednit tzv. odlehle vzory typicky zatížené chybami měření. Pro uvažované funkce uchazečka odhadla i tzv. míru

aproximace (rates of approximation). V práci je dále navržen a podrobně analyzován originální pravděpodobnostní algoritmus použitelný i pro prořezávání neuronových sítí s extrémně velkým počtem skrytých neuronů.

Většinu původních výsledků uváděných v práci autorka publikovala na mezinárodních recenzovaných konferencích, novější výsledky pak formou technických zpráv vydaných ÚI AV ČR. Text práce je napsaný dobrou angličtinou (drobnější výhrady se týkají např. překlepů – ř. 14 sh. na str. 39, ř. 13 zd. na str. 89, ř. 5 zd. na str. 93, použitého značení – ř. 5 sh. na str. 94, ř. 7 zd. na str. 104, resp. neúplných citací – ř. 8 sh. na str. 111, ř. 15 sh. na str. 114). Vlastní práce pak zahrnuje detailní popis, analýzu a vzájemné porovnání studovaných modelů. K předkládané práci mám následující otázky a připomínky:

1. Vlastnosti analyzovaných modelů by bylo vhodné porovnat s dalšími známými typy neuronových sítí používanými pro aproximaci funkcí – např. s modelem tzv. ELM-sítí (Extreme Learning Machine) G.-B. Huanga.

Další otázky: Jak se projeví v práci navrhovaná regularizační technika založená na pravděpodobnostním přístupu na generalizačních schopnostech sítě (VC-dimenzi vytvářené sítě)? Lze tuto metodu kombinovat s dalšími algoritmy, např. principem učení s nápovědou (learning with hints) Abu-Mostafy?

2. Předkládaná práce je zaměřená na studium aproximačních schopností dopředných neuronových sítí s jednou skrytou vrstvou neuronů. Je však známo, že sítě se dvěma anebo větším počtem skrytých vrstev se při aproximaci chovají odlišně.

Otázka: Jak se projeví větší počet vrstev na kvalitě a použitelnosti získaných výsledků (např. míra aproximace, počet resp. vlastnosti prořezávaných neuronů apod.)?

3. K práci, bohužel, není přiložena implementace, resp. zdrojové kódy testovaných modelů a metod. Práce nicméně obsahuje stručnou dokumentaci některých doposud provedených experimentů.

Otázky: Jak byla použita metoda křížové validace při určování parametrů (ř. 15 zd. na str. 85)? Jakých výsledků se autorce podařilo získat při dalším testování metod (např. při prořezávání skrytých neuronů)?

### Závěr

Předkládaná práce podle mého názoru splňuje požadavky kladené na disertační práci. Uchazečka se v práci zabývá vysoce aktuální problematikou efektivního zpracování velkého množství vícerozměrných dat s využitím masivně paralelních distribuovaných výpočetních modelů založených na paradigmatu umělých neuronových sítí. Stanovené cíle se disertantce podařilo splnit.

Vědecký přínos předložené práce spatřuji především v důkladné teoretické analýze aproximačních schopností dopředných neuronových sítí s jednou skrytou vrstvou a v návrhu originálního pravděpodobnostního algoritmu použitelného i pro prořezávání neuronových sítí s extrémně velkým počtem skrytých neuronů. Výsledky doposud provedených testů naznačují použitelnost navržených modelů i při řešení poměrně obtížných úloh z praxe (predikce průtoku na řece Ploučnici). Experimentální ověření vlastností dalších vyvinutých technik disertantka provádí ve spolupráci s RNDr. Petrou Kudovou - Vidnerovou, Ph.D. z Ústavu Informatiky AV ČR.

Předloženou práci proto doporučuji k obhajobě a po jejím úspěšném obhájení doporučuji udělit Mgr. Terezii Šámalové titul Ph.D.

V Praze, 9. 8. 2008



Doc. RNDr. Iveta Mrázová, CSc.

KTIML MFF UK, Univerzita Karlova v Praze  
Malostranské náměstí 25, 118 00 Praha 1