

# VYJÁDRĚNÍ ŠKOLITELE K DISERTAČNÍ PRÁCI

*Téma práce:*       **Artificial Neural Networks and Their Usage  
for Knowledge Extraction**

*Autorka práce:*   **RNDr. Zuzana Petříčková, KTIML MFF UK, Praha**

---

## Aktuálnost tématu práce

V oboru počítačových věd už dnes umělé neuronové sítě představují zavedené paradigma. Zejména vrstevnaté neuronové sítě jsou známé díky své schopnosti dobře zobecňovat a odhalit v datech složité nelineární závislosti. Na druhou stranu má tento model tendenci vytvářet poměrně složitou vnitřní strukturu, která se následně jen velmi obtížně interpretuje. Praktické využití těchto modelů při řešení reálných úloh vyžadujících zpracování velkého množství typicky vícerozměrných dat proto bylo donedávna poměrně omezené. Kromě vysoké dimenze vstupních dat a pomalého učení totiž bývá problematické i jejich zatížení šumem. Použité algoritmy pak mohou i pro relativně podobná data poskytovat různé výsledky.

V souvislosti s širším využitím umělých neuronových sítí pro řešení náročných úloh z praxe tak nabývá na významu studium problematiky týkající se rychlosti učení a volby vhodné architektury použitého modelu, jeho generalizačních schopností a interpretace extrahovaných znalostí. Předkládaná disertační práce RNDr. Zuzany Petříčkové patří do zmíněné perspektivní oblasti výzkumu. Předmětem práce je návrh nových postupů pro učení vrstevnatých neuronových sítí typu zpětného šíření a jejich využití při extrakci znalostí a pravidel z velkých datových sad.

## Cíl práce

Při řešení reálných úloh pomocí umělých neuronových sítí je kladen důraz především na rychlost procesu učení a schopnost zobecnit správně extrahované znalosti. Ukazuje se však, že jedním z klíčových prvků pro adekvátní generalizaci může být i transparentní architektura trénované sítě. Vytvoření transparentní struktury sítě totiž výrazně usnadňuje následné prořezání redundantních neuronů - skrytých i vstupních. Nalezení optimálního modelu konzistentního s uvažovanými daty přitom představuje jednu ze základních úloh strojového učení.

Cílem předkládané práce proto bylo navrhnout efektivní algoritmus pro učení vrstevnatých neuronových sítí typu zpětného šíření, který by podporoval snadnou interpretaci extrahovaných znalostí. Vyvinutý algoritmus měl přitom být robustní jak vzhledem k volbě nastavitelných volných parametrů, tak také k drobným odchylkám ve vstupních datech. Vlastnosti nových přístupů měla uchazečka důkladně otestovat s ohledem na jejich případné využití při zpracování velkých vícerozměrných dat, např. z oblasti ekonomie.

## Metody zpracování

Úvodní část předkládané práce je věnována problematice vrstevnatých neuronových sítí a jejich učení. Další analyzované přístupy se týkají výběru příznaků, technik pro optimalizaci struktury vytvářených neuronových sítí, zvyšování jejich generalizačních schopností a strategií pro vytváření transparentní struktury. Na základě podrobné analýzy řešené problematiky uchazečka navrhla obecnou metodologii pro učení vrstevnatých neuronových sítí s vynucovanou interní reprezentací a oslabovanou citlivostí modelu k drobným odchylkám ve vstupních datech. Redundantní vstupní a skryté neurony jsou prořezávány pomocí technik založených na citlivostní analýze a interní reprezentaci znalostí.

K rozsáhlému testování původních i referenčních technik použila disertantka vlastní knihovnu implementovanou v integrovaném prostředí MATLAB. Systém využívá i některé funkce aplikační

knihovny Neural Network Toolbox a je k práci připojen na přiloženém CD. Použitelnost navrhovaného řešení uchazečka ověřila jak na uměle vygenerovaných datech, tak také na reálných datech ze Světové banky. Vlastní text práce je napsaný kultivovaným jazykem a svědčí o profesionální orientaci autorky ve studované oblasti.

### Dosažené výsledky

Práce přináší několik nových strategií pro efektivnější učení vrstevnatých neuronových sítí typu zpětného šíření. Nové techniky vycházejí z metody škálovaných konjugovaných gradientů a jejich výhodou je, že je lze navzájem kombinovat. Jednotlivé algoritmy učení (škálované konjugované gradienty s vynucovanou interní reprezentací i škálované konjugované gradienty s exaktně, resp. aproximativně oslabovanou citlivostí) se totiž liší v podstatě jen použitou regularizační technikou. Předností nově navržených metod pak je zejména:

- lepší generalizační schopnost vytvářených sítí a jejich nižší citlivost k šumu v datech,
- rychlé učení robustní vzhledem k nastavitelným parametrům,
- vytváření jednodušší a transparentnější interní struktury sítě (především díky tzv. kondenzované interní reprezentaci vynucované během učení a potlačované citlivosti sítí vzhledem k drobným odchylkám na vstupních datech),
- robustnost učených sítí vzhledem k počáteční volbě topologie a redundantním vstupním příznakům (díky prořezávání skrytých a vstupních neuronů)
- a snazší interpretace znalostí extrahovaných učenou neuronovou sítí.

Vzhledem k relativně nízkým výpočetním nárokům se pro praktické účely jeví jako nejvhodnější technika škálovaných konjugovaných gradientů s aproximativně oslabovanou citlivostí. Naučené sítě totiž dobře zobecňují a podporují i efektivní prořezávání. Získané poznatky mají ovšem nezanedbatelný potenciál i při zpracování velkých dat s využitím programovatelného hardwaru (např. GPU), a to především v souvislosti s vývojem nových modelů určených pro analýzu multimediálních dat (např. hlubokých neuronových sítí).

### Závěr

Předkládaná práce splňuje všechny požadavky kladené na disertační práci. Uchazečka se v práci zabývá vysoce aktuální problematikou redukce dimenzionality vstupních dat, optimalizace architektury umělých neuronových sítí a interpretace extrahovaných znalostí. Dosažené výsledky vznikly v rámci několika grantových projektů a jsou podpořeny celou řadou publikací zejména z významných mezinárodních konferencí, kde několikrát získaly ocenění typu "Best Paper Award." Stanovené cíle se tedy disertantce podařilo beze zbytku splnit.

Vědecký přínos předložené práce spatřuji především v návrhu několika originálních postupů důležitých pro efektivní učení robustních modelů vrstevnatých neuronových sítí, a to především s ohledem na snadnou interpretaci extrahovaných znalostí. Vyvinuté metody vycházejí z moderních principů používaných v oblasti strojového učení, umělých neuronových sítí i dobývání znalostí a tyto poznatky podstatným způsobem dále rozvíjejí. Předloženou práci proto doporučuji k obhajobě a po jejím úspěšném obhájení doporučuji udělit RNDr. Zuzaně Petříčkové titul Ph.D.

V Praze, 5. 8. 2015

doc. RNDr. Iveta Mrázová, CSc.  
KTIML MFF UK, Praha