

## POSUDEK OPONENTA BAKALÁŘSKÉ PRÁCE

Název:	Optimalizační algoritmy pro trénování neuronových sítí
Autor:	Jiří Szotkowski

### SHRNUTÍ OBSAHU PRÁCE

Práce se zabývá popisem a trénováním neuronových sítí z matematického hlediska. Neuronová síť je představena jako zobrazení, které vektoru vstupů (podnětů) a aktuálním parametrům sítě přiřazuje výstupní vektor (aktivity neuronů). Je formulován problém trénování neuronové sítě, ve kterém je snahou určit parametry sítě tak, aby výstupy neuronové sítě odpovídaly na daném datasetu co nejlépe očekávaných výstupům. To lze vyjádřit v podobě minimalizace tzv. ztrátové funkce. V práci je detailně popsán algoritmus, který umožňuje spočítat gradient ztrátové funkce. Na hledání lokálního minima ztrátové funkce tak lze použít vhodné spádové metody. Autor dále popisuje optimalizační algoritmy založené na modifikacích metody největšího spádu, které se běžně k trénování neuronových sítí používají. Vysvětluje také myšlenku použití stochastických verzí těchto algoritmů. Konečně, v závěrečné kapitole jsou numericky testovány klasické i stochastické verze algoritmů, na akademických i praktických příkladech.

### CELKOVÉ HODNOCENÍ PRÁCE

**Práce** je rešerší známých výsledků. Vlastním **příspěvkem** autora je velmi pěkný a detailní popis dané problematiky nejen z pohledu matematického, ale i pedagogického. V práci je ukázán postup od matematické formulace problému až po jeho numerické řešení, včetně náznaků směrů dalšího možného pokračování ve výzkumu. Matematický **text** je logicky členěn a korektně zformulován. Původní **zdroje** jsou správně a v dostatečné míře citovány. Práce obsahuje minimum tiskových chyb.

### ZÁVĚR

Práci považuji za velmi kvalitní a doporučuji ji uznat jako bakalářskou práci.

V Praze, 28. května 2024

doc. RNDr. Petr Tichý, Ph.D.  
KNM, MFF UK

## OTÁZKY K DISKUSI

1. Uvažované optimalizační spádové metody jsou většinou modifikace metody největšího spádu. Tyto modifikace se snaží volit následující směrový vektor jako vhodnou lineární kombinaci směru největšího spádu a předchozího směrového vektoru. To nápadně připomíná například nelineární metodu sdružených gradientů, která rovněž volí nový směrový vektor z lineárního obalu zmíněných vektorů. Proč není nelineární metoda sdružených gradientů vhodná k minimalizaci ztrátové funkce?
2. Je známo něco bližšího o matematických vlastnostech ztrátové funkce? (konvexita, existence sedlových bodů, atd.)